UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

LUIZ FILIPE DE MACEDO BARTOLETI

FILOGEOGRAFIA DAS ESPÉCIES DE *Nephila* (ARANEAE: ARANEIDAE) DA AMÉRICA DO SUL

PHYLOGEOGRAPHY OF *Nephila* (ARANEAE: ARANEIDAE) SPECIES FROM SOUTH AMERICA

Campinas

2017

LUIZ FILIPE DE MACEDO BARTOLETI

FILOGEOGRAFIA DAS ESPÉCIES DE *Nephila* (ARANEAE: ARANEIDAE) DA AMÉRICA DO SUL

PHYLOGEOGRAPHY OF *Nephila* (ARANEAE: ARANEIDAE) SPECIES FROM SOUTH AMERICA

Tese apresentada ao
Instituto de Biologia da Universidade
Estadual de Campinas como parte
dos requisitos exigidos para a
obtenção do título de Doutor
em Genética e Biologia Molecular,
na área de Genética Animal e Evolução

Thesis presented to the
Institute of Biology of the University
of Campinas in partial fulfillment
of the requirements for the degree of
Doctor in Genetics and Molecular Biology,
in the area of Animal Genetics and Evolution

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELO ALUNO
LUIZ FILIPE DE MACEDO BARTOLETI,
E ORIENTADA PELA PROFA.
DRA. VERA NISAKA SOLFERINI

Orientadora: Profa. Dra. Vera Nisaka Solferini

Campinas

2017

Agência(s) de fomento e nº(s) de processo(s): CNPq, 141040/2013-0; FAPESP, 2012/02526-7
ORCID: http://orcid.org/0000-0002-0724-5057

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

B285f

Bartoleti, Luiz Filipe de Macedo, 1988-
Filogeografia das espécies de *Nephila* (Araneae: Araneidae) da América do Sul / Luiz Filipe de Macedo Bartoleti. – Campinas, SP : [s.n.], 2017.

Orientador: Vera Nisaka Solferini.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Filogeografia. 2. Aranha. 3. Florestas tropicais. 4. Biodiversidade - América Latina. 5. Diagonal seca. I. Solferini, Vera Nisaka,1957-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Phylogeography of *Nephila* (Araneae: Araneidae) species from South America
Palavras-chave em inglês:
Phylogeography
Spiders
Tropical forests
Biodiversity - Latin America
Dry Diagonal
Área de concentração: Genética Animal e Evolução
Titulação: Doutor em Genética e Biologia Molecular
Banca examinadora:
Vera Nisaka Solferini [Orientador]
André Victor Lucci Freitas
Samantha Koehler
Fernando de Faria Franco
Clarisse Palma da Silva
Data de defesa: 16-08-2017
Programa de Pós-Graduação: Genética e Biologia Molecular

Campinas, 16 de Agosto de 2017

COMISSÃO EXAMINADORA

Profa. Dra. Vera Nisaka Solferini (Presidente)

Prof. Dr. André Victor Lucci Freitas

Profa. Dra. Samantha Koehler

Prof. Dr. Fernando de Faria Franco

Profa. Dra. Clarisse Palma da Silva

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A todos os agentes de mudança...

AGRADECIMENTOS

Agradeço à Vera por novamente confiar em meu trabalho e realizar a orientação do meu Doutorado. Por toda a ajuda e dedicação ao longo dos últimos quatro anos, que foram essenciais para a minha formação como cientista.

Aos membros da banca, pelas críticas construtivas e fundamentadas, que contribuíram grandemente para a melhoria do trabalho realizado.

À minha família, em especial meus pais, Ademir e Sueli, e irmãos, Thiago e Vitor. Pelo apoio e compreensão, carinho e cuidado.

Aos amigos do LDG, e em especial os que tanto me ajudaram no doutorado: Bia, Elen, Fernanda, Jair, Livia, Natalia e Thadeu. Pelas discussões de trabalho e, mais importante, pelos momentos passados fora dele.

A todos os ajudantes de campo e de laboratório, que tornaram esse trabalho possível.

Aos grandes amigos que fiz em Campinas durante os muitos e intermináveis anos de UNICAMP. Em especial àqueles dos quais me aproximei nos últimos dois anos, e que me ajudaram das mais diversas formas.

Ao Interbio, agora mais do que nunca.

Ao bandejão (e todos os funcionários dele!), por permitir que eu e outros milhares de alunos tivéssemos condições de nos manter na Universidade com alimentação de qualidade.

Ao Programa de Pós Graduação em Genética e Biologia Molecular, ao Instituto de Biologia e à UNICAMP, por proporcionarem a infraestrutura e os recursos necessários para a formação de jovens cientistas. A todos os professores e funcionários que me ajudaram de alguma forma.

À CAPES, à FAPESP e ao CNPq por fornecerem os fundos necessários para que esse e outros projetos fossem realizados.

RESUMO

A região Neotropical é reconhecida por sua grande diversidade biológica; no entanto, os mecanismos que levaram à essa alta diversificação ainda não são bem compreendidos. Neste trabalho, realizamos o estudo filogeográfico das duas espécies Neotropicais do gênero *Nephila* a fim de contribuir para a compreensão dos diferentes processos que influenciaram a diversificação de espécies Neotropicais, e apresentamos os resultados em dois manuscritos. No primeiro, "Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest", buscamos elucidar como eventos geológicos e climáticos afetaram a diversificação de espécies adaptadas a florestas úmidas. Utilizando um marcador mitocondrial (COI) e dois nucleares (H3a e ITS2), detectamos cinco linhagens geograficamente informativas. A linhagem que se encontra a oeste da Cordilheira Central dos Andes colombianos divergiu das demais no final do Mioceno/Plioceno, concomitantemente ao soerguimento final dessa Cordilheira. As demais linhagens datam do Pleistoceno, o que sugere uma diversificação devido aos eventos climáticos do Quaternário. A distribuição geográfica das linhagens, e o teste de modelos feito por Approximate Bayesian Computation (ABC) indicam um isolamento prévio entre os biomas, com posterior contato secundário entre as linhagens no último máximo glacial (LGM – há 21 mil anos), provavelmente devido à mudança na distribuição dos biomas de acordo com as flutuações climáticas do Quaternário. A diagonal seca parece ter atuado como uma fonte de migrantes para Amazônia e Mata Atlântica, e nossos dados corroboram uma via de conexão entre as florestas pelo meio do Cerrado. No segundo manuscrito, "Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal" utilizamos populações de *Nephila sexpunctata* para realizar inferências sobre a história recente da diagonal seca. Utilizamos dois marcadores mitocondriais (COI e ND1-L1-16S) e um nuclear (CHP2), e detectamos uma linhagem associada ao Cerrado e uma ao Chaco, com alguma mistura. A modelagem de paleodistribuição realizada indicou uma expansão do nicho da espécie a partir do Holoceno Médio (cerca de 6 mil anos atrás), o que concordou com a recente expansão demográfica inferida com os dados genéticos. O teste de modelos feito com ABC indicou que a espécie passou por um severo gargalo populacional no LGM (que se reflete na baixa variabilidade genética encontrada), com uma expansão demográfica posterior ao glacial, contrariando o que seria esperado pelo Modelo dos Refúgios Pleistocênicos, evidenciando que o glacial pode ter apresentado condições climáticas severas demais até mesmo para espécies adaptadas a ambientes abertos. Nosso trabalho evidenciou que os estudos com espécies de

ampla distribuição são fundamentais para a melhor compreensão da história evolutiva de regiões altamente diversas. Os dois trabalhos sugerem que a região Neotropical possui uma história bastante complexa, e que os padrões evolutivos são fortemente influenciados por características biológicas da espécie. Estudos com organismos menos representados como as aranhas devem ser encorajados pois são capazes de revelar padrões interessantes da história biogeográfica da região Neotropical.

ABSTRACT

The Neotropical region is known by its great biological diversity; however, the mechanisms that drove this diversification are yet not well understood. In this work, we performed the phylogeographical study of the two Neotropical *Nephila* species aiming to contribute to the comprehension of different processes that influenced the diversification of Neotropical species, and we present the results in two manuscripts. In the first one, "Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest", we sought to elucidate how geological and climate events affected the diversification of rainforest-adapted species. Using one mitochondrial (COI) and two nuclear (H3a and ITS2) molecular markers, we detected five geographically informative lineages. The lineage present westwards of the Colombian Central Andean Cordillera has diverged from the others in the late Miocene/Pliocene, concomitantly with the final uplift of this Cordillera. The other lineages emerged in the Pleistocene, what suggests a diversification guided by the Quaternary climate events. The geographical distribution of the lineages and the Approximate Bayesian Computation (ABC) model testing indicated a previous isolation among biomes, with posterior secondary contact among lineages in the Last Glacial Maximum (LGM – 21kya), probably due to the changes in biome distribution according to the Quaternary climate fluctuations. The Dry Diagonal seems to have provided migrants for both Amazon and the Atlantic Forest, and our data corroborate a connection route between the rainforests through the center of Cerrado. In the second manuscript, "Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal" we used *Nephila sexpunctata* populations to make inferences on the recent history of the Dry Diagonal. We used two mitochondrial (COI and ND1-L1-16S) and one nuclear (CHP2) molecular markers, and detected one lineage associated with Cerrado and one with Chaco, with some admixture. The palaeodistribution modelling indicated a niche expansion for this species in the middle Holocene (around 6 kya), what agrees with the recent demographic expansion inferred from the genetic data. The ABC model testing indicated a severe bottleneck in the LGM (which agrees with the low genetic diversity in the species), with a post-glacial demographic expansion, against what was expected by the Pleistocene Refugia Model, evidencing that the glacial may have presented harsh climate conditions even for species adapted to open environments. Our work evidenced that studies with widely distributed species are fundamental to a better comprehension of the evolutionary history of highly diverse regions. Both works suggest that the Neotropical region has a complex history, and that evolutionary

patterns are strongly affected by biological peculiarities from each species. Studies with less represented taxa as spiders must be encouraged as they may unveil interesting patterns about the biogeographical history of the Neotropical region.

SUMÁRIO

INTRODUÇÃO GERAL

A região Neotropical

Dentre as regiões biogeográficas, os Neotrópicos – que se estendem do sul da Flórida até a Argentina – se destacam por apresentarem elevados índices de diversidade. Vários biomas nesta região são considerados prioridades mundiais para conservação, devido às altas taxas de endemismo e avançado estágio de degradação (Myers *et al*, 2000). A região constitui um mosaico de diferentes tipos de vegetação, que inclui florestas tropicais úmidas (como Amazônia e Mata Atlântica), formações secas (principalmente na chamada "Diagonal Seca", que engloba a Caatinga, no nordeste brasileiro; o Cerrado, na região central do Brasil; e o Chaco, no norte da Argentina, Paraguai e Bolívia), florestas montanas (nas encostas de grandes cadeias de montanhas, como os Andes), vegetações subtropicais etc. (Ab'Saber, 1977; Carvalho & Almeida, 2011).

A paisagem do continente sul-americano foi intensamente transformada nos últimos milhões de anos. Anteriormente, em períodos quentes e úmidos como o Eoceno (56 a 33,9 milhões de anos atrás), as formações vegetais úmidas se estendiam ao longo de todo o continente, constituindo uma grande área contínua de florestas tropicais na América do Sul (Morley, 2000). A partir de então, eventos em diferentes escalas afetaram a paisagem da região, influenciando diretamente na diversificação dos organismos que a habitavam. Eventos geológicos, datados principalmente do Neógeno, parecem ter tido papel importante nesse processo; dentre eles, o soerguimento da cadeia montanhosa dos Andes (entre o Mioceno e o Plioceno), além de funcionar como evento vicariante, provocou alterações drásticas no clima, regime de chuvas e sistemas hídricos (Hoorn *et al*, 2010; Rull, 2011). O fechamento do Istmo do Panamá, também estimado entre o final do Mioceno e o Plioceno (Winston *et al*, 2017), promoveu o intercâmbio entre as biotas da América do Sul e das Américas Central e do Norte, além de alterar a dinâmica das correntes marítimas e, por conseguinte, afetar o clima da região. As mudanças locais no clima ocasionadas por esses eventos, aliadas a uma diminuição global da precipitação e da temperatura, propiciaram um aumento na ocorrência das gramíneas C_4, principalmente na região central do Brasil (Simon *et al*, 2009; Potter & Szatmari, 2009). A Diagonal Seca emergiu a partir dessa expansão das paisagens savânicas, e provocou uma marcante disjunção no grande cinturão de florestas tropicais úmidas, dando origem à Amazônia e à Mata Atlântica. Após a sua formação, a Diagonal Seca também sofreu a ação de eventos geológicos que influenciaram a diversificação nessa região, como o soerguimento do Planalto

Brasileiro no Mioceno (com o concomitante rebaixamento do Chaco e outras regiões baixas – Silva, 1995; Carvalho *et al*, 2013), e recorrentes transgressões marítimas no Chaco ao longo do Mioceno (Ruskin *et al*, 2011).

Além dos eventos geológicos, eventos climáticos do Quaternário também são frequentemente elencados como potenciais promotores de diversificação na região Neotropical. Desde a proposição da Teoria dos Refúgios Pleistocênicos (Haffer, 1969; Vanzolini & Williams, 1981), a alternância entre períodos quentes e úmidos com períodos frios e secos, bastante pronunciada nessa época, foi estabelecida como promotora de isolamento entre diferentes fragmentos de um mesmo bioma. As populações isoladas em cada fragmento, então, evoluiriam independentemente (Carnaval & Moritz, 2008), o que poderia acarretar em diferenciação de linhagens e até especiação. Com o aumento do número de trabalhos na América do Sul que mostram que houve intensa diversificação de linhagens no Quaternário (revisado por Turchetto-Zolet *et al*, 2013), a influência de eventos climáticos na diversificação de organismos da região parece ter ganhado nova força. Às flutuações climáticas, também, são atribuídas as alterações na distribuição geográfica dos biomas ao longo do Pleistoceno, que teriam acompanhado as mudanças na temperatura e pluviosidade (Sobral-Souza *et al*, 2015). Dessa maneira, a história dos biomas Neotropicais ao longo do Pleistoceno teria sido dinâmica, e conexões entre Amazônia e Mata Atlântica, mesmo depois do estabelecimento da Diagonal Seca, são hipotetizadas por diversos autores (e.g. Por, 1992; Costa, 2003; Batalha-Filho *et al*, 2013; Prates *et al*, 2016).

Embora um conjunto de trabalhos tenha proposto hipóteses acerca da diversificação da região Neotropical, o número de estudos que efetivamente testou essas hipóteses ainda é pequeno, tendo em vista a complexidade dos padrões observados (Beheregaray, 2008; Turchetto-Zolet *et al*, 2013).

A Filogeografia aplicada ao estudo da diversificação Neotropical

A filogeografia surgiu ao final da década de 1980, numa tentativa de aplicar os métodos e conhecimentos filogenéticos ao nível microevolutivo, estudando espécies ou pares de espécies (Avise *et al*, 1987). Usando como base a teoria da coalescência, a filogeografia permite inferir eventos demográficos do passado que tenham influenciado a atual distribuição de linhagens ou grupos gênicos dentro de espécies (Avise, 2009). Dessa maneira, a extrapolação desses resultados possibilita procurar eventos comuns que tenham afetado diretamente um grupo de espécies de distribuição similar, e até mesmo que tenham afetado todo um bioma.

Apesar de a filogeografia ter suas bases em análises com DNA mitocondrial, os estudos filogeográficos tem cada vez mais buscado usar marcadores nucleares, uma vez que as tecnologias atuais de sequenciamento permitem a obtenção relativamente rápida de um grande número de marcadores moleculares (Garrick *et al*, 2015) e, em consequência, inferências mais robustas dos parâmetros demográficos. Aliado a isso, o grande avanço nas metodologias de análise de dados e a disseminação de técnicas como a modelagem de paleodistribuição (revisado em Lima-Ribeiro & Diniz-Filho, 2012), conjuntamente com o conhecimento da história geológica e climática da região, permitiram que novas hipóteses *a priori* fossem criadas. O surgimento da filogeografia estatística (Knowles & Maddison, 2002) foi de suma importância para aumentar a robustez das inferências históricas, pois as diversas hipóteses evolutivas passaram a ser testadas e o grau de incerteza das análises passou a ser estimado. A integração de todos esses métodos permite uma avaliação estatística de modelos que levam em conta as peculiaridades biológicas do organismo estudado (em oposição à mera comparação de modelos genéricos) e que consideram a estocasticidade inerente dos processos evolutivos que moldaram a distribuição da variabilidade genética em cada espécie (Knowles *et al*, 2007).

Na região Neotropical, os estudos filogeográficos ainda são poucos em comparação a outros continentes, embora em constante ascensão. Revisados em 2013 por Turchetto-Zolet *et al*, os trabalhos revelaram um forte viés em estudos com vertebrados (69% dos trabalhos), ou organismos que ocorram em apenas uma região climática (89%), ou em apenas um bioma (47%), e utilizando apenas o DNA mitocondrial (58%). Assim, há uma grande necessidade de que diferentes organismos, como os invertebrados, também sejam utilizados nesse tipo de estudo, uma vez que podem elucidar diferentes aspectos da história evolutiva dos biomas. Além disso, a utilização de organismos amplamente distribuídos pode contribuir para uma visualização mais acurada dos processos evolutivos de biomas que tem uma história intricada. Por fim, a utilização do DNA mitocondrial conjuntamente com o nuclear permite avaliar diferentes aspectos da história evolutiva das espécies.

As aranhas como modelos de estudos filogeográficos

As aranhas (Arachnida: Araneae) constituem uma das mais diversas ordens dentre os invertebrados terrestres, com 46.762 espécies descritas atualmente, divididas em 112 famílias (World Spider Catalog, 2017). Essa grande diversidade se reflete em características biológicas de diferentes grupos. Além da grande variação quanto à tolerância ambiental, há uma grande variação nos modos de dispersão: algumas espécies dispersam pouco tanto na fase juvenil

quanto na adulta, e geralmente constroem as teias próximas às da mãe (e.g. *Aglaoctenus lagotis* - Santos & Brescovit, 2001), enquanto outras são capazes de praticar "balonismo", ou dispersão aérea, na fase jovem (Bell *et al*, 2005) – quando os juvenis são carregados pelo vento através de fios de teia e podem se deslocar por diversos quilômetros (e.g. *Nephila pilipes* – Lee *et al*, 2015).

A ordem Araneae tem sido utilizada como modelo filogeográfico em todos os continentes (e.g. África – Kuntner & Agnarsson, 2011; Ásia e Oceania – Su *et al*, 2007; Europa – Bidegaray-Batista *et al*, 2016; América do Norte – Crews & Gillespie, 2014). Na América do Sul, entretanto, o número de trabalhos filogeográficos envolvendo esse grupo ainda é reduzido. Os estudos de Magalhães *et al*, 2014 e Peres *et al*, 2015, foram os primeiros a utilizar aranhas como modelos na região. Ambos trabalhos implementaram uma abordagem estatística e contrastaram com sucesso diferentes modelos demográficos, conseguindo realizar inferências sobre o impacto das mudanças climáticas do Pleistoceno em diferentes partes do continente. Esses trabalhos evidenciam que a realização de estudos com aranhas amplamente distribuídas nos Neotrópicos pode ser essencial para a melhor compreensão de eventos biogeográficos dessa região.

As espécies de *Nephila* da América do Sul

A classificação taxonômica de *Nephila* e seus gêneros próximos (*Clitaetra*, *Herennia*, *Nephilengys*, *Nephilingis*) tem sido amplamente debatida na última década. Em 2006, Kuntner estabeleceu a família Nephilidae através de análises etológicas e de morfologia, composta pelos gêneros acima citados (à exceção de *Nephilingis*, que foi separado de *Nephilengys* por Kuntner *et al*, 2013) anteriormente pertencentes à família Tetragnathidae. Entretanto, desde a elevação de Nephilidae à categoria de família, a relação desse clado com as famílias próximas nunca foi devidamente suportada. As principais hipóteses indicavam que Nephilidae poderia ser o grupo irmão de Araneidae (Kuntner, 2006; Su *et al*, 2011) ou de Araneidae+Tetragnathidae (Álvarez-Padilla *et al*, 2009). Dimitrov *et al*, 2016, utilizando seis marcadores moleculares e dados de arquitetura da teia, transferiu os cinco gêneros nefilídeos para a família Araneidae, recriando a subfamília Nephilinae. Nós optamos por seguir essa classificação mais recente nos trabalhos desenvolvidos.

O gênero *Nephila* é pantropical e possui 23 espécies reconhecidas. Nas Américas, há ocorrência de apenas duas espécies: *Nephila clavipes*, distribuída ao longo de toda a região Neotropical, dos Estados Unidos à Argentina; e *Nephila sexpunctata*, restrita à porção central

do Brasil, Paraguai, Bolívia e norte da Argentina. Espécies do gênero tem sido utilizadas com sucesso como modelo para estudos sobre a evolução da vida em grupo (Hodge & Uetz, 1992), comportamento sexual (Quiñones-Lebrón *et al*, 2016), realocação e estrutura das teias (Blamires *et al*, 2010), além de trabalhos com biogeografia e filogeografia em outros continentes (Su *et al*, 2007, 2011). No entanto, o conhecimento de aspectos biogeográficos e de variabilidade genética das espécies de *Nephila* na América do Sul, sobretudo de *N. sexpunctata*, é praticamente inexistente.

OBJETIVOS

Esta tese teve como objetivo principal compreender os processos históricos e evolutivos que afetaram a diversificação de organismos da região Neotropical, por meio da análise filogeográfica das duas espécies do gênero *Nephila* (Araneae: Araneidae) da América do Sul.

Os dois manuscritos que compõe essa tese abordam diferentes aspectos da diversificação na região: eventos geológicos e climáticos que influenciaram a história evolutiva de organismos que ocorrem em florestas úmidas, com ênfase em conexões Pleistocênicas entre Amazônia e Mata Atlântica são discutidos no Manuscrito I, a partir dos padrões filogeográficos de *Nephila clavipes*; e a história demográfica recente de organismos da diagonal de vegetações secas, com ênfase no Cerrado e Chaco, é o foco do Manuscrito II, com *Nephila sexpunctata*. As duas abordagens, em conjunto, apresentam um panorama mais preciso dos processos responsáveis pela diversificação recente na região.

Dentre os objetivos específicos dos trabalhos, estão:

- Quantificar a variabilidade genética de cada espécie e avaliar como ela está estruturada;
- Verificar a presença de linhagens genéticas geograficamente informativas;
- Estimar o tempo de divergência entre tais linhagens, relacionando a eventos climáticos ou geológicos característicos da região;
- Inferir eventos demográficos nas espécies estudadas;
- Testar hipóteses que auxiliem no melhor entendimento da evolução dos organismos da região Neotropical.

MATERIAL E MÉTODOS

Organismos de estudo

As espécies do gênero *Nephila* tecem grandes teias orbiculares, que podem chegar a até 1,5m de diâmetro (Kuntner, 2006). Os fios possuem coloração dourada, responsável pelo nome popular de "*golden orb web spiders*".

Nephila clavipes (Linnaeus, 1767) é a espécie do gênero mais amplamente distribuída nos Neotrópicos, ocorrendo desde os Estados Unidos até o norte da Argentina. É abundante principalmente nas formações florestais úmidas, como Amazônia e Mata Atlântica, mas também pode ocorrer em enclaves florestais e matas de galerias em vegetações mais secas, como o Cerrado. Apresenta ciclo de vida anual (Moore, 1977), em que as fêmeas tecem as teias perto de estradas, trilhas ou corpos d'água, em microclimas bastante úmidos (Rypstra, 1985; Moore, 1977; Robinson & Mirick, 1971). A espécie apresenta um acentuado diformismo sexual, sendo as fêmeas muito maiores que os machos; estes, na vida adulta, não constroem teia própria e vivem nas teias das fêmeas, onde se alimentam e reproduzem (Moore, 1977).

Figura 1: Fêmeas de *Nephila clavipes* (esquerda) e *Nephila sexpunctata* (direita).

Nephila sexpunctata (Giebel, 1867) possui uma distribuição mais reduzida, estando restrita à região sudoeste da Diagonal Seca, composta pelo Cerrado, Chaco e outras formações xéricas adjacentes, englobando a região central do Brasil, Paraguai, norte da Argentina e sul da Bolívia. Detalhes da biologia da espécie são profundamente desconhecidos, visto que trabalhos anteriores com a espécie são quase inexistentes. A espécie apresenta características comuns a

todas espécies do gênero como o ciclo de vida anual, o marcante dimorfismo sexual e as grandes teias douradas tecidas em árvores ou camadas arbustivas altas (observações pessoais).

Ambas espécies, assim como as demais componentes do gênero, são consideradas boas dispersoras (Kuntner & Agnarsson, 2011). Diversos trabalhos com espécies asiáticas demonstraram ausência de estruturação genética entre indivíduos separados por distâncias de mais de seis mil quilômetros ou por grandes cadeias montanhosas (Tso *et al*, 2002; Lee *et al*, 2004; Su *et al*, 2007; Su *et al*, 2011), o que sugere boa manutenção de fluxo gênico. Recentemente, foi empiricamente constatado o comportamento de balonismo em juvenis de *N. pilipes* (Lee *et al*, 2015), o que reforça que as espécies do gênero devem ser capazes de dispersar por longas distâncias. Na fase adulta, os machos parecem ser responsáveis por uma maior manutenção de fluxo gênico, uma vez que são capazes de se mover de teia em teia de acordo com o sucesso alcançado no acasalamento (Rittschof, 2010). A menor movimentação das fêmeas pode gerar padrões de alta estruturação do DNA mitocondrial, como observado por Cooper *et al*, (2011) e Croucher *et al*, (2011).

Apesar de serem as duas únicas espécies do gênero na região Neotropical, *N. sexpunctata* e *N. clavipes* não são espécies irmãs, e representam distintas linhagens do gênero *Nephila* que teriam chegado ao continente americano através de eventos independentes de dispersão a longa distância a partir da África, Ásia ou Oceania no Mioceno/Plioceno. De acordo com a filogenia proposta por Kuntner *et al*, 2013, baseada em marcadores mitocondriais e nucleares, morfologia e comportamento, *Nephila clavipes* possui como espécies mais próximas as africanas *N. senegalensis*, *N. fenestrata*, *N. inaurata*, *N. turneri*, *N. komaci* e *N. sumptuosa*. Já *Nephila sexpunctata* possui divergência mais recente com espécies da Ásia e Oceania, como *N. edulis*, *N. plumipes*, *N. antipodiana* e *N. clavata*.

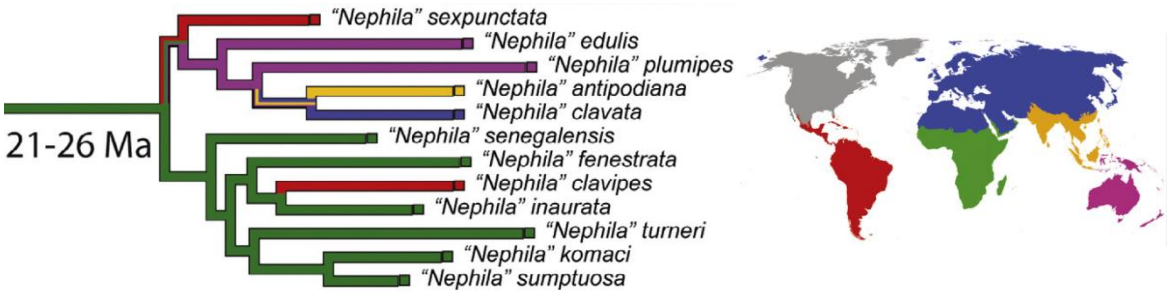


Figura 2: Filogenia parcial para o gênero *Nephila* reconstruída por Kuntner *et al*, 2013, indicando os principais clados dentro do gênero e o continente em que cada espécie ocorre.

Desenho amostral

As coletas para as duas espécies foram planejadas de modo que amostrássemos amplamente a distribuição conhecida, a partir de dados de museus e de banco de dados como specieslink (http://splink.cria.org.br/) e gbif (http://www.gbif.org/). Para *N. sexpunctata*, foram amostrados pontos no Cerrado brasileiro, além de pontos no Chaco seco e Chaco úmido na Argentina. Para *N. clavipes*, foram amostrados pontos nas partes Norte e Sul da Mata Atlântica, nas florestas de galeria no Cerrado, na Amazônia e em quatro localidades na Colômbia (sendo duas a Leste da Cordilheira Central – Amazônia Colombiana e Valle del Magdalena; e duas a Oeste da Cordilheira Central – Valle del Cauca e próximo à costa do Oceano Pacífico). As coletas foram realizadas entre 2009 e 2016, preferencialmente nos meses iniciais do ano (entre Janeiro e Maio), pois nos meses mais chuvosos há maior abundância de indivíduos adultos. O detalhamento do desenho amostral encontra-se na Figura 2.

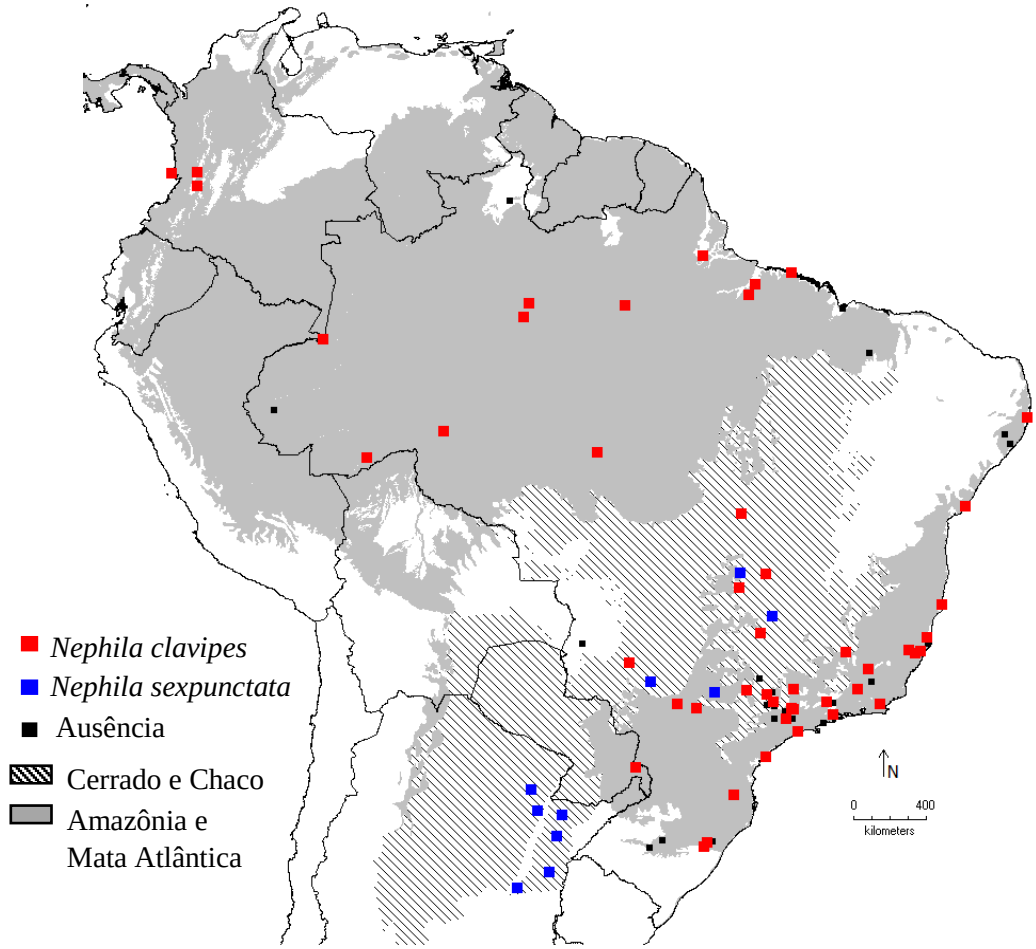


Figura 3: Desenho amostral mostrando as localidades em que foram coletados indivíduos de *N. clavipes* (em vermelho) e *N. sexpunctata* (em azul).

Tabela 1: Localidades de coleta, com número de indivíduos utilizados nas análises, para *N. clavipes* e *N. sexpunctata*.

	N. clavipes		***N. sexpunctata***	
	Localidade	**N**	**Localidade**	**N**
1	Alcobaça /BA	8	Argentina – Basail	8
2	Alta Floresta/MT	8	Argentina – Mercedes	8
3	Aracruz/ES	8	Argentina – Paraná	3
4	Belém/PA	8	Argentina – PN del Chaco	7
5	Belo Horizonte/MG	8	Argentina – Santa Maria	8
6	Bragança/PA	8	Argentina – Villa Federal	8
7	Brasília/DF	8	Campo Grande/MS	4
8	Campo Grande/MS	1	Catalão/GO	3
9	Catalão/GO	1	Pirenópolis/GO	8
10	Colombia – Amazon	1	Tupã/SP	8
11	Colombia – V. del Magdalena	6	Total	65
12	Colombia – Pacific	3		
13	Colombia – V. del Cauca	2		
14	Florestópolis/PR	8		
15	Foz do Iguaçu/PR	8		
16	Goiânia/GO	3		
17	Gramado/RS	8		
18	Ibirama/SC	8		
19	Ibitinga/SP	8		
20	Ilha do Cardoso/SP	8		
21	Itaguaçu/ES	3		
22	Juiz de Fora/MG	8		
23	Jundiaí/SP	8		
24	Linhares/ES	8		
25	Macapá/AP	3		
26	Manaus/AM	8		
27	Mata de São João/BA	8		
28	Moju/PA	8		
29	Monte Alegre do Sul/SP	8		
30	Paraty/RJ	7		
31	Passa Quatro/MG	8		
32	Pinhalzinho/SP	8		
33	Pirenópolis/GO	8		
34	Poços de Caldas/MG	8		
35	Porangatu/GO	8		
36	Porto Velho/RO	6		
37	Recife/PE	3		
38	Rio Branco/AC	8		
39	Rio Claro/SP	8		
40	Rio Preto do Eva/AM	1		
41	Santa Teresa/ES	1		
42	Santarém/PA	8		
43	Santos/SP	8		
44	São Carlos/SP	8		
45	Sapiranga/RS	8		
46	Silva Jardim/RS	7		
47	Teodoro Sampaio	8		
48	Uberlândia/MG	8		
49	Viçosa/MG	8		
	Total	320		

Marcadores moleculares e análises

Nos dois manuscritos presentes neste trabalho, obtivemos sucesso ao usar conjuntamente marcadores mitocondriais e nucleares, que revelam em diferentes profundidades aspectos da história evolutiva das espécies estudadas (Zhang & Hewitt, 2003). Utilizamos o marcador mitocondrial Citocromo Oxidase *c* subunidade I (COI), que vem sendo amplamente utilizado em estudos filogeográficos, inclusive com diferentes grupos de aranhas Neotropicais (e.g. Magalhaes *et al*, 2015; Peres *et al*, 2015). No manuscrito I, com *N. clavipes*, utilizamos ainda duas regiões nucleares: o gene codificante da subunidade *a* da histona H3 (H3a), além do espaçador interno transcrito II (ITS2), de DNA ribossômico. Ambas regiões já haviam sido amplificadas com sucesso para *N. clavipes* (Kuntner *et al*, 2013) e utilizadas em estudos filogeográficos com aranhas (H3a - Magalhaes *et al*, 2015; ITS2 - Peres *et al*, 2015), em que se provaram bastante informativos. Já no manuscrito II, com *N. sexpunctata*, utilizamos adicionalmente: a região de DNA mitocondrial composta pelo gene da subunidade grande de rRNA 16S, o RNA transportador de Leucina e o gene da enzina NADH-ubiquinona oxidoredutase (16S-L1-NAD1); e uma região intrônica de DNA nuclear prospectada em nosso laboratório, denominada CHP2.

Em ambos manuscritos realizamos análises de diversidade e estruturação genética, fizemos inferências filogenéticas Bayesianas multilocus com estimativas de tempo de divergência entre as principais linhagens, construímos redes de haplótipos para todos os marcadores e realizamos inferências demográficas. No manuscrito I, com *N. clavipes*, também inferimos a área ancestral das principais linhagens mitocondriais a fim de entender possíveis mudanças na distribuição de cada uma delas. No manuscrito II, com *N. sexpunctata*, fizemos uma modelagem de paleodistribuição, a fim de reconstituir a área de distribuição da espécie ao longo do último ciclo glacial (Último Interglacial, ~120 mil anos atrás; Último Máximo Glacial, ~21 mil anos atrás; Holoceno Médio, ~6 mil anos atrás; e presente) e, assim, dispor de mais informações relevantes para a formulação de hipóteses demográficas. Por fim, para as duas espécies também implementamos uma abordagem de teste de modelos, em que contrastamos as diferentes hipóteses demográficas para cada espécie e selecionamos a que melhor explica os nossos dados genéticos através de Approximate Bayesian Computation (ABC - Beaumont, 2010; Csilléry *et al*, 2010; Sunnaker *et al*, 2013).

MANUSCRITO I

Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest

(Manuscrito submetido ao periódico *Journal of Biogeography*)

ORIGINAL ARTICLE:

Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest

Luiz Filipe de Macedo Bartoleti[1] (LFMB); Elen Arroyo Peres[2] (EAP); Fernanda von Hertwig Mascarenhas Fontes[1] (FVHMF); Márcio José da Silva[3] (MJS); Vera Nisaka Solferini[1]* (VNS)

[1]Department of Genetics, Evolution and Bioagents, Institute of Biology, University of Campinas, Campinas, São Paulo, Brazil

[2]Department of Zoology, Institute of Biosciences, University of São Paulo, São Paulo, São Paulo, Brazil

[3]Center of Molecular Biology and Genetic Engineering, University of Campinas, Campinas, São Paulo, Brazil

*Correspondence: Vera N. Solferini, Laboratory of Genetic Diversity, Department of Genetics, Evolution and Bioagents, Institute of Biology, University of Campinas, Rua Monteiro Lobato 255, 13083-862, Campinas, São Paulo, Brazil.

E-mail: veras@unicamp.br

Running head: Phylogeography of *Nephila clavipes*

Word count: 6989

Word count (abstract): 249

Number of pages for figures: 3-4.

ABSTRACT

Aim The South-American rain forests share a common biogeographic origin and have an interconnected history that includes the drier biomes between them. It is not clear the degree of isolation promoted by the establishment of the Neotropical Dry Diagonal and if connections between the forests have occurred after this event. We sought to elucidate aspects on these biomes' biogeography through a phylogeographical study of *Nephila clavipes*, a rain forest-dweller spider.

Location South America.

Methods We collected 320 individuals from 49 sites through the species' range and sequenced one mitochondrial and two nuclear DNA regions. We performed analyses on genetic diversity and structure, demography, phylogenetic inferences and estimation of divergence times, inferred ancestral areas and applied a model-based approach in order to test competitive hypotheses on the species' evolution.

Results An ancient split, coincident with the Central Cordillera final uplift, segregated a Trans-Andean Colombian lineage from the others. Four geographically informative lineages occur in Brazil, with a Pleistocene divergence. The biogeographical analysis indicated that some lineages have expanded their ranges towards other biomes along Pleistocene.

Main Conclusions The distribution of lineages, as well as the model testing results, corroborate a rain forest connection through the central part of Cerrado taking place at the Last Glacial Maximum. Our data corroborates that geological and climate events have affected Neotropical diversification, and reinforces that studying widely distributed species is primordial to unveil evolutionary patterns in regions with an entangled and intricate history.

Keywords Amazon, Atlantic Forest, Cerrado, Pleistocene climate fluctuations, Neotropics, phylogeography

INTRODUCTION

The Neotropics harbor two large tropical forests, Amazon (AM) and Atlantic Forest (AF), acknowledged by high species richness and endemism. These forests, once continuous during the wet and warmer Eocene, are currently separated by a large corridor of drier environments, called Dry Diagonal (DD) – formed by Caatinga, Cerrado and Chaco –, which emerged after events of cooling and dryness along Oligocene and Miocene (Morley, 2000); however, the underlying biogeographic events are still not completely understood.

Some of the major diversification events in South America are attributed to Neogene geologic/orogenic activities, such as the final uplift of the Andean Cordilleras (Hoorn *et al*, 2010), estimated around the Miocene/Pliocene. The formation of this massive mountain range caused great impact in the region's climate, hydric systems and rainfall regime (Hoorn *et al*, 2010), promoting vicariance and diversification. Alternatively, Quaternary climate fluctuations are increasingly been cited as a strong source of intraspecific variation in the Neotropics (Turchetto-Zolet *et al.*, 2013).

Several studies show that DD settlement has affected the Neotropical biota evolution, isolating previously widespread forest taxa also promoting lineage diversification (Por, 1992; Costa, 2003). DD has possibly acted as a barrier to these organisms' dispersion, and populations in each side have thus evolved independently. Some groups (such as amphibians and birds) present sister species in AM and AF with divergence times estimated in the Miocene or earlier (Batalha-Filho *et al*, 2013; Fouquet *et al*, 2014), reinforcing this idea. Other works, however, suggest putative rain forests connections after DD emergence, since sister lineages in each biome diverged in the Pliocene or Pleistocene (Costa, 2003; Batalha-Filho *et al*, 2013; Prates *et al*, 2016; Peres *et al*, 2017). Such connections have probably occurred due to shifts on biome distribution according to climate cycles (Sobral-Souza *et al*, 2015). Moreover, present-day patches of humid vegetation within the DD form a net of gallery forests (Oliveira-Filho & Ratter, 1995), which preserve suitable habitats for rainforest-dwellers, possibly maintaining current connections between AM and AF.

Phylogeography has been increasingly employed to study the evolution of megadiverse regions, as the Neotropics. However, most researches focus on already well studied groups such as birds, anurans and plants (reviewed by Turchetto-Zolet *et al*, 2013), while few have used less represented invertebrate taxa, despite they being abundant, highly diverse and suitable to unveil evolutionary patterns worldwide (Moritz *et al*, 2001). Spiders, notably, have received more attention in other continents, but few studies in the Neotropics have addressed

phylogeographical questions using this group (Magalhaes *et al*, 2014; Peres *et al*, 2015; Bartoleti *et al*, 2017; Peres *et al*, 2017). Spiders exhibit a wide range of environmental resilience (e.g. tolerance to dryness or low temperatures) and dispersal abilities (some species lack ballooning behavior – when spiderlings are pulled by the wind through silk threads –, while others have been reported to float for large distances); in a context of dynamic connections between rain forests, these disparities may affect differently how species can handle harsh conditions in an inhospitable matrix and how far individuals can get when connections are established.

Nephila clavipes (Linnaeus, 1767) is a widespread Neotropical araneid, abundant in rain forests and in humid patches within open environments as Cerrado. As evidenced for other *Nephila* species (Lee *et al*, 2015), *N. clavipes* is supposed to successfully perform aerial dispersal. The species likely arrived in the Neotropics in the Miocene/Pliocene through a long-distance dispersal event from Africa (Kuntner *et al*, 2013), and diversified in the continent. Given its wide distribution, ancient history in South America and biological peculiarities, *N. clavipes* is a suitable candidate for a phylogeographical study focusing on the history of Neotropical forests.

In this work we investigated the historical biogeography of Neotropical rain forests through a phylogeographical study of *N. clavipes*. We used one mitochondrial and two nuclear markers and coupled population genetic analyses, divergence times estimation and a model-testing framework to clarify how historical events affected diversification in this region. We proposed to test specific non-exclusive hypotheses concerning Neotropical evolution: (i) geological events from the Neogene shaped the diversification; (ii) patterns of diversification were caused by climate events along Quaternary. In addition, we sought to test whether connections among Brazilian biomes have occurred and when they took place. Therefore, we tested if these connections (a) happened due to Pleistocene climate fluctuations, mainly in the Last Glacial Maximum (LGM ~21 ka) or (b) these connections are still active in the present through the gallery forests within DD. Testing these hypotheses may contribute to a better understanding of the dynamics experienced by Neotropical forests that promoted diversification in the continent.

MATERIALS AND METHODS

Sample collection

The sampling design covered most of the species' distribution in South America, through AM (66 individuals from 10 localities), DD (69 individuals from 11 localities), AF (173 individuals from 24 localities), including ecotones and transitional areas. We also sampled individuals from four localities in Colombia [in the Northern portion of the Andes, which are split into three Cordilleras [Eastern – EC, Central – CC and Western – WC (Cooper *et al*, 1995)]: i) Amazonian Colombia, in the border with Brazil (one individual); ii) Magdalena Valley, in the eastern slopes of CC (six individuals); iii) Cauca Valley, between CC and WC (two individuals); iv) Pacific Colombia, close to the Pacific Ocean (three individuals) (Fig. 1). For some analyses, we grouped populations according to the biome where they occur, following the classification of IBGE (Brazilian Institute of Geography and Statistics): AM, DD, Northern Atlantic Forest (NAF) and Southern Atlantic Forest (SAF). We used the Doce River (~20ºS) to distinguish NAF and SAF as this region is frequently invoked as a place of species turnover in this biome (Carnaval *et al*, 2014) and it also marks a transition between two different climate regimes (Cheng *et al*, 2013). In locations where the official classification was dubious (i.e. transitional areas), we used field observations to assign biomes. Colombian samples were also separated in groups: Amazonian Colombia and Magdalena Valley, in the eastern side of CC, were included in the group AM (see Results); Cauca Valley and Pacific Colombia, in the western side of CC, constituted a group of Trans-Andean populations (TAC).

DNA extraction, amplification, sequencing and alignment

Genomic DNA was extracted from 1-8 legs using the Wizard Genomic DNA Purification kit (Promega) following the manufacturer's protocol. Species were identified by female genitalia or male palps. Vouchers were deposited in the Coleção Científica de Aracnídeos e Miriápodes of Instituto Butantan (São Paulo, Brazil).

We used the cytochrome *c* Oxidase subunit I (COI) as mitochondrial marker, and two nuclear regions: the coding portion of the Histone 3 subunit a (H3a) and the Internal Transcribed Spacer 2 (ITS2). For the mitochondrial marker, we obtained sequences for all sampled individuals, while for the nuclear markers we amplified a representative subsample (123 individuals for H3a and 99 for ITS2 – Table 1, Fig. S1). Amplifications were obtained using the sets of primers: LCO1490 and HCO2198 (Folmer *et al*, 1994) for COI; H3aF2 and H3aR2 (Colgan *et al*, 1998) for H3a; and 5.8S and ITS4 (White *et al*, 1990) for ITS2.

PCR conditions varied for each marker: a denaturation step of 3-4' at 94ºC, 30 cycles at 94ºC for 45", 46-61ºC for 45-60" and 72ºC for 1-2', and a final extension step at 72ºC for 3-

10'. The standard PCR reaction consisted of 1μL of template DNA, 0.2μL of 5U Taq DNA Polymerase (Thermo-Fisher), 2.5μL of 10X KCl Buffer (Thermo-Fisher), 2.5-3.5μL of 25mM $MgCl_2$ (Thermo-Fisher), 0.5-1μL of 10mM dNTP mix (GE Healthcare), 0.4-0.5μL of each forward and reverse primer 10μM, and ddH_2O to the volume of 25μL. Sequences were obtained in an automatic capillary sequencer ABI PRISM 3700 DNA Analyzer (Applied Biosystems). We performed the alignment in MEGA 7.0 (Kumar *et al,* 2016), using the *Muscle* algorithm (Edgar, 2004), visually inspecting each sequence.

Genetic diversity, population structure and haplotype reconstruction

We calculated standard genetic diversity indices for each marker, assessed population structure within species through F_{ST} and performed an AMOVA using biomes as geographic groups in ARLEQUIN 3.5 (Excoffier & Lisher, 2010). We used DnaSP (Librado & Rozas, 2009) to infer gametic phases from the nuclear markers, with a minimum posterior probability threshold of 0.9. We also used BAPS 6.0 (Corander *et al*, 2008) to determine the most likely number of genetic clusters within our sequences (testing from 1 to 20). We performed separate analyses for the mitochondrial and nuclear datasets, using two replicates to ensure the reliability of the results.

We built median-joining networks (Bandelt *et al*, 1999) on POPART (Leigh & Bryant, 2015) to assess relationships among haplotypes.

Phylogenetic inferences and divergence times

Substitution models that best fitted each dataset were defined according to the Akaike Information Criterion (AIC) on jMODELTEST 2.0 (Darriba *et al*, 2012). We conducted a multilocus species tree Bayesian phylogenetic analysis (*BEAST) in BEAST v1.8.0 (Heled & Drummond, 2010). We used lognormal relaxed clocks, as the test for the molecular clock performed on MEGA 7.0 rejected the strict clock hypothesis. For the mitochondrial dataset, we used a substitution rate of 0.0112 substitutions/site/million years, based on a geological calibration for the Dysderidae family in the Mediterranean basin (Bidegaray-Batista & Arnedo, 2011) and considered an adequate estimate of mitochondrial evolution for nephilids based on a fossil calibration for this group (Kuntner *et al*, 2013). For the nuclear data, we used an estimate of 0.0013 substitutions/site/million years (Bidegaray-Batista & Arnedo, 2011) for H3a substitution rate (as there is no rate information available for ITS2 evolution on spiders). We chose as outgroup the African species *N. inaurata*, the species closest to *N. clavipes* according

to the phylogeny proposed by Kuntner *et al*, 2013 (GenBank accession numbers: KC849085.1, KC849044.1 and JF835932 for COI, H3a, and ITS2 respectively).

We grouped the sequences in five mitochondrial lineages inferred by BAPS (see Results) and performed 1 billion MCMC simulations (storing 10,000 trees). We checked the convergence with a stationary distribution in TRACER v1.5 (Rambaut *et al*, 2014), discarded the first 10% of the trees as burn-in and determined the best MCC species and gene trees with TREEANNOTATOR v1.8.0, visualized and edited in FIGTREE v1.3.1 (Rambaut, 2012). To access uncertainty among different gene trees, we built a density tree in DENSITREE (Bouckaert & Heled, 2014).

Demographic patterns

We inferred demographic patterns through neutrality tests (Tajima, 1989; Fu, 1997) and analyses of mismatch distribution performed in ARLEQUIN 3.5, separately for each marker and for each mitochondrial group (see Results). We also conducted Extended Bayesian Skyline Plot analyses (EBSP) on BEAST v1.8.0 to detect demographic fluctuations over time in this species. We performed analyses for the whole dataset, as well as for each mitochondrial group, with the same nucleotide substitution rates as in the *BEAST analysis. We followed the tutorial for the analysis found in the BEAST website.

Reconstruction of ancestral areas

We considered each biome (AM, DD, NAF, SAF and TAC) as discrete units and performed the reconstruction of *N. clavipes* ancestral areas using a Bayesian Stochastic Search Variable Selection (BSSVS) on BEAST v1.8.0. The analysis was performed with the mtDNA dataset using only one sequence per haplotype per population (154 individuals). We followed the steps on the BEAST tutorial for this analysis; nucleotide substitution model and COI substitution rate were the same as the used in the phylogenetic tree analysis, and we applied the lognormal relaxed molecular clock. The coalescent tree prior was set to a constant population size, and inference of ancestral areas was performed using a symmetric substitution model under a strict clock. We performed a run with 30 million generations (sampling every 3,000 generations) and checked stationarity on TRACER v1.5. Steps to generate a MCC tree were similar to those used in the *BEAST analysis.

Model Testing

We used an Approximate Bayesian Computation framework (ABC) to perform the model testing. We separated the individuals in geographic groups (TAC, AM, DD and AF) and constructed demographic models that represent biogeographical hypothesis concerning the evolutionary history of the Neotropical region (Fig. 2). We grouped similar models (which represent alternatives for the same hypothesis) in scenarios and performed the model testing hierarchically (as in Peres *et al*, 2015; Bartoleti *et al*, 2017), to ultimate select which hypothesis was the best overall. Scenario 1 represents a single panmictic population; Scenario 2 represents a Neogene split between TAC and the other samples, coincident with the CC final uplift (see Discussion), and a large panmictic population in the Brazilian samples. For scenarios 3-5 we considered that, after the Neogene TAC split, climate fluctuations promoted diversification among Brazilian biomes during Quaternary. For Scenario 3 we considered that AM samples diverged from the panmictic population constituted by AF+DD in the Pleistocene; in Scenario 4, AF and DD had the most recent divergence in the Pleistocene; in Scenario 5, all Brazilian biomes diverged due to a fragmentation event at the same time in the Pleistocene. In scenarios 3-5, we tested for i) current multidirectional migration; ii) current non-uniform migration; iii) LGM multidirectional migration; iv) LGM non-uniform migration. We performed the analyses jointly for all markers. More details in the models, scenarios and the methodology can be assessed in the Appendix S1.

RESULTS

Genetic diversity

We obtained a 594bp COI segment, and HKY+G was the best fitting substitution model for this dataset. We found a high number of haplotypes (63), high haplotype and nucleotide diversity (0.937 and 0.0111, respectively – Table 2). Nucleotide diversity decreased substantially within biomes while haplotype diversity remained high for most of them. DD and SAF had higher diversity indices, attributable to the high admixture between different genetic groups (see *Genetic Structure* section).

The H3a dataset presented sequences of 243bp and had K2+G+I as the best substitution model. Diversity levels were lower than those from COI (24 haplotypes, Hd = 0.7428 and π = 0.0089 – Table 2) and were similar among biomes, except for TAC populations which presented a unique haplotype for all individuals.

The ITS2 dataset spanned 211bp and GTR+I was the most fitting substitution model. While haplotype diversity and number of haplotypes were lower than H3a (Hd = 0.6132, H =

7), nucleotide diversity was equivalent (π = 0.009 – Table 2). As for H3a, haplotype diversity was similar in the overall dataset and within each biome; for the nucleotide diversity, levels were lower within biomes. TAC populations presented a single haplotype for all its individuals. Details on diversity within each population are presented on Tables S1-S2.

Genetic Structure

The BAPS analysis identified five mitochondrial groups (Fig. 1): Group I is composed by all NAF samples, most of SAF and DD; Group II encompasses AM and Colombian samples in the eastern side of CC; Group III is composed by samples widespread from southern AM, through DD, reaching the southernmost SAF, but occurs mostly in the SAF-DD transition zone; Group IV occur predominantly in DD, but some individuals were found in southern AM and in SAF; Group V is found exclusively in the western side of CC in Colombia. Individuals from Group V are separated from the others by a minimum of 32 mutational steps (5.8% of divergence; average = 1.1%), while the other groups bare a closer relationship. F_{ST} for COI was very high (0.751) and decreased when calculated within each biome (Table 2). Pairwise F_{ST} values were mostly high, and TAC populations were responsible for the maximum values (Table S3). AMOVA using biomes as groups showed that nearly 50% of the genetic variability is organized among them (Table 3); when TAC populations were excluded from the analysis, the among-groups variation was reduced to ~38%.

H3a sequences presented lower genetic structure (F_{ST} = 0.243), but the value for ITS2 was also high (F_{ST} = 0.716). These datasets did not recover the groups evidenced by the mitochondrial marker: for H3a, all TAC individuals presented the same exclusive haplotype, separated from the others by only one mutational step. For ITS2, TAC individuals also presented an exclusive haplotype separated from the others by 11 mutations (5.2% of divergence; average = 0.2%). For both nuclear markers, the remaining haplotypes were mostly shared among biomes. BAPS analysis for the nuclear dataset also evidenced that TAC represents a group distinct from the other biomes, showing a weaker structure in Brazilian samples. AMOVA results highlighted this trend, with about 10% of H3a variability found among biomes, while the majority was present within populations. For ITS2, most of the variability was present among biomes (70%); when TAC populations were not considered in the AMOVA, the among-groups variation strongly decreased in both datasets (non-significant values - Table 3). Pairwise F_{ST} values for the nuclear markers were low, but also presented TAC comparisons as the maximum values (Tables S4 and S5).

Phylogenetic inferences and divergence times

Divergence between *N. clavipes* and *N. inaurata* was estimated in the Pliocene/Miocene, around 5.6 Ma (95%HPD = 2.97-9.56, Fig. 3). TAC divergence from Brazilian and eastern Colombia samples was estimated around 3.39 Ma (95%HPD = 1.78-5.64, Fig. 3). All lineages that occur in Brazil diverged in the Pleistocene, in a short period of time (~ 0.3 Ma, from 0.342 to 0.045 Ma, Fig. 3). Groups indicated by BAPS are strongly supported in the mitochondrial tree (Fig. 3), but the relationships among groups are dubious as posterior probabilities are low. NAF individuals are disposed intertwined with SAF and DD individuals in Group I. SAF and DD individuals occur in groups I, III and IV, while AM individuals occur almost exclusively in Group II, with few individuals in Groups III and IV. TAC individuals belong exclusively to Group V. In the nuclear trees, the only well-defined group is TAC, while the clustering of the other individuals was not geographically informative (data not shown). The multilocus species tree (Fig. 3, left) show the great uncertainty among the gene trees, mainly among the four Brazilian lineages.

Demographic patterns

Neutrality tests and mismatch distribution analyses for the mitochondrial marker presented signals of recent demographic expansion in the overall dataset, Groups I and II, and the biomes AM and SAF. Nuclear markers did not present significant results (Table 2, Fig. S2).

The multilocus EBSP analyses detected demographic changes for Groups I and IV, with evidence of recent expansion around 200 ka and 10 ka, respectively (Fig. 4).

Reconstruction of ancestral areas

We could not infer with high posterior probability an ancestral location for all the individuals. Brazilian samples seem to coalesce somewhere between DD (43%) and SAF (35%), while TAC samples differentiated in western Colombia (95%). In Brazil, the ancestral area inferred for each lineage was the biome where they occur more frequently: SAF for Group I (65%), AM for Group II (99%) and DD for Groups III and IV (50 and 79%, respectively – Fig. 3, Table S6).

Model Testing

When performed among scenarios, the model selection held model 10 as the best fit for our empiric data, with a strong posterior probability (0.7493). Therefore, the best demographic

scenario for our data comprises a split between TAC and the other samples around the time of CC final uplift (~10-6 Ma), a more ancient split between AM and AF+DD, and a final and more recent split between AF and DD. It also corroborates an exchange of migrants mainly between DD and each forest along the LGM (Table 4).

DISCUSSION

Our genetic data and coalescence-based model testing revealed a scenario in which Neogene geological events and Quaternary climate fluctuations are responsible for diversification in the Neotropics. Moreover, our data support the occurrence of connections between AM and AF during LGM, possibly caused by shifts in biomes' distribution due to climate changes in this period. DD seems to play an important role in this process, providing migrants for both forests. AF and DD present a closer relationship, possibly due to the exchange of individuals represented by a zone of secondary contact between the states of São Paulo and Minas Gerais (Fig. 1a, lower detail), as described for other organisms in the same region (D'Horta *et al*, 2011; Dantas *et al*, 2015).

Phylogeographical patterns of Nephila clavipes

The molecular markers presented adequate levels of genetic variation to unveil interesting patterns of diversification in this species. These levels are compatible with previous studies using *Nephila* in other continents (Su *et al*, 2007; Su *et al*, 2011), but strongly contrast with what was found for the other Neotropical species in the genus, *Nephila sexpunctata* (Bartoleti *et al*, 2017). However, low diversity levels for *N. sexpunctata* were associated with demographic events that decreased diversity, which are not likely to have occurred for *N. clavipes*. For both species, the diversification patterns in the mitochondrial dataset were not entirely reflected in neither nuclear marker. The discrepancies between nuclear and mitochondrial markers may be caused by differences in effective population size and evolutionary rates, incomplete lineage sorting, mtDNA introgression and/or differential dispersal between sexes (Toews & Brelsford, 2012).

Even though *N. clavipes* and *N. sexpunctata* are the only *Nephila* species in the Neotropics, they represent distinct lineages of the genus with independent evolutionary histories, as their ancestors supposedly arrived in South America through distinct long-distance dispersal events from Africa, Asia or Oceania (Kuntner *et al*, 2013). The estimated divergence between *N. clavipes* and the African *N. inaurata* in the Miocene/Pliocene supports the idea of

dispersal to the Neotropics (Kuntner *et al*, 2013), previously corroborated for *N. sexpunctata* (Bartoleti *et al*, 2017). These results endorse that Trans-Atlantic exchange of biota was common in the Neogene (Renner, 2004; Christenhusz & Chase, 2013). The forest-dweller *Nephila* lineage from which *N. clavipes* would emerge arrived in South America after DD settlement and established itself; since then, experienced several events that shaped its genetic structure.

The Northern Andes as a driver for diversification

We detected a conspicuous distinction between Trans-Andean populations and the other samples through our data. All markers presented only exclusive haplotypes for TAC populations, showing great genetic divergence from the other haplotypes, and the high indices of genetic structure seems to be due to the large difference between TAC and the other populations.

The Northern Andes, region where our Colombian populations were sampled, are characterized by three Cordilleras with distinct geological histories: CC and WC had a more ancient final uplift in the Miocene, while EC reached its current height more recently, in the Pliocene (Gregory-Wodzicki, 2000). This asynchronous uplifting seems to have affected the genetic structure of vertebrate species, isolating different phylogroups in WC+CC and EC (e.g. Valderrama *et al*, 2014); however, the diversification time is not always congruent among studies. *Rheobates* frogs from CC and EC seem to have diverged in the Miocene (Muñoz-Ortiz *et al*, 2014). For the more vagile *Metallura* hummingbirds, divergence between WC+CC and EC dated to Pleistocene (Benham *et al*, 2015).

The divergence between *N. clavipes* populations from both sides of the CC (Pacific and Cauca Valley in the west, and Magdalena Valley and Amazon in the east) was estimated in the Miocene/Pliocene, consistent with CC final uplift (as reinforced by the ABC), suggesting that it may be a barrier to dispersion in this species. This finding differs from a previous study with *Nephila pilipes* in southeastern Asia, in which the authors found that the Central Mountain Range (CMR) in Taiwan does not prevent gene flow between populations from different sides of the mountains (Lee *et al*, 2004). However, the CC in Colombia has an average height of almost 4,000 meters and several peaks above that height, while CRM is considerably lower, what may explain the diminished gene flow in this case.

Despite the high elevations of EC, it does not seem to be a barrier for *N. clavipes*. The population from the Magdalena Valley is genetically similar to the others from Colombian and Brazilian Amazon. In a similar manner, some *Rheobates* frogs from the eastern and western

slopes of EC are closely related, what suggests that individuals may have crossed EC after the diversification of this lineage, dated to Pleistocene. The authors inferred that this crossing might have happened through the Paso de Andalucía, a depression in the southern portion of EC as low as 2,000m, possibly facilitated by Pleistocene climate fluctuations (Muñoz-Ortiz *et al*, 2014). *N. clavipes* might have dispersed similarly, as the lineage that reached the Magdalena Valley (Group II) dates back to ~170 ka, in the Pleistocene.

Pleistocene diversification in Brazilian biomes

Mitochondrial data for *N. clavipes* revealed the existence of four geographically associated lineages in Brazilian biomes (Fig. 1), with divergence times estimated in the Pleistocene. Therefore, the recent coalescence of our data indicates that DD settlement (~10 Ma) was not a vicariant event for the species, suggesting some degree of connectivity between AM and AF up to the Quaternary. Divergences occurred mainly in the Middle-Pleistocene (~300 ka), marked by intense climate oscillations with increased amplitude (Head & Gibbard, 2005). In this scenario, shifts in biome distribution (Sobral-Souza *et al*, 2015) might have isolated *N. clavipes* populations in disjoint stable areas that would allow populations to evolve independently (Carnaval & Moritz, 2008), originating the mitochondrial lineages. Moreover, three of the Brazilian lineages (Groups I, III and IV) seem to have emerged somewhere between DD and SAF in south-central Brazil, a region severely affected by climate fluctuations during Pleistocene (Morley, 2000).

Even though these lineages occur together in some sampling locations, they present great genetic divergence and different ancestral areas, what suggests moments of isolation during the Quaternary with recent secondary contact caused by the species' high dispersal ability. ABC results also supported differentiation among biomes with posterior migration among previously isolated populations. Recent works with widely distributed spiders in the Neotropics have also found geographically informative lineages with a Pleistocene divergence (Peres *et al*, 2017), suggesting that climate fluctuations may be responsible by the recent diversification.

Demographic events for forest-dwellers Neotropical species have been discussed since the Refugia Theory (Haffer, 1969; Vanzolini & Williams, 1981), which suggested that wet forests would have retracted during glacial periods and expanded during interglacials. Even though some organisms present this expected pattern (e.g. De Ré *et al*, 2014), recent works have challenged this concept as several organisms display population stability (e.g. Batalha-

Filho *et al*, 2012) or pre-LGM expansion (e.g. Peres *et al*, 2015). For *N. clavipes*, analyses evidenced demographic expansion in the lineages that preferably inhabit rain forests (Groups I and II, AF and AM – Fig. 4, Table 2). No dataset presented evidence of population retraction. The EBSP estimated a subtle demographic expansion for Group I (the most associated with AF) around 200 ka, before the LGM. This suggests that this lineage expanded or at least remained stable during the last glacial cycle, in opposition to the expected by Refugia Theory. The possibility of AF range expansion towards the continental shelf during the LGM (Leite *et al*, 2016), together with *N. clavipes* capability of building webs in several different habitats (Moore, 1977) and handling a wide range of environmental conditions, might help explaining the demographic stability found throughout our data, as responses to climate fluctuations depend heavily on the species' ecological and environmental tolerances (Prates *et al*, 2016).

Putative connections between AM and AF

There is substantial evidence supporting the existence of AM-AF Pleistocene connections including palaeopalynogical, geological and phylogeographical data (Behling *et al*, 2000; Auler *et al*, 2004; Batalha-Filho *et al*, 2013; Peres *et al*, 2017). Hypothesized connections routes include i) a corridor through the north of Caatinga, linking eastern-AM to NAF; ii) a path linking western-AM and SAF through Chaco and Cerrado; iii) a connection through central Brazil, as forest fragments can be currently found in Cerrado (Por, 1992; Costa, 2003; Batalha-Filho *et al*, 2013).

The distribution of *N. clavipes* lineages is not restricted to biome boundaries, with three lineages (Groups I, III and IV) occurring in more than one biome. According to the ancestral area reconstruction, Group I was likely originated in SAF and then expanded to DD and NAF, while Groups III and IV were formed in the DD and reached SAF and AM (Figs. 1 and 3, Table S6). These data corroborate the "Central Cerrado" route (Por, 1992; Oliveira-Filho & Ratter, 1995) following a corridor through the central part of Cerrado, reaching both AM and AF (Fig. 1). The model testing also supported this hypothesis, showing that previously isolated biomes have exchanged migrants during LGM, with DD being the main source for both forests. Climate simulations have shown that central Brazil was more humid during LGM due to an expansion of SAF climate towards DD (Sobral-Souza *et al*, 2015), which may have created intermittent corridors connecting SAF-DD-AM. This pattern was presumably recurrent along the Quaternary as glacial cycles repeatedly changed climate conditions in central Brazil (Cheng *et al*, 2013). As a species with high dispersal ability and environmental resilience, capable of

building webs at the borders of forest fragments (Moore, 1977), *N. clavipes* might have expanded its range through these corridors when climate conditions were adequate.

Studies with several organisms (Batalha-Filho *et al.*, 2013; Rodrigues *et al.*, 2014; Prates *et al*, 2016) , including widespread rain forest-dweller spiders (Peres *et al*, 2017), support a close relationship between eastern-AM and NAF. This connection would have been established through the expansion of AM climate towards NAF (Sobral-Souza *et al*, 2015), and is corroborated by the existence of the "brejos de altitude", rain forest fragments in the Caatinga highlands. However, *N. clavipes* did not presented haplotype sharing nor trans-occurrence between these biomes. All haplotypes present in NAF were also present in SAF. Studies encompassing NAF and SAF recurrently report the existence of different lineages in each counterpart, what has been associated with a region of climate turnover near of the Doce River (Cheng *et al*, 2013; Sobral-Souza *et al*, 2015; Peres *et al*, 2017). The great similarity between NAF and SAF haplotypes in this study indicates that *N. clavipes* populations in the AF maintain a great level of gene flow, therefore no phylogeographical breaks could be inferred for this biome.

CONCLUSIONS

Our data support that both geological (Andean CC uplift) and climate (Pleistocene fluctuations) events were responsible for Neotropical diversification at different time frames (Neogene x Quaternary). We also elucidated that DD settlement was not a vicariant event for *N. clavipes*, as the coalescence of Brazilian samples took place in the Pleistocene, suggesting Quaternary connectivity between AF and AM. ABC results supported the hypothesis that AM-AF connections happened during LGM with DD as a major source of migrants, and our data seem to fit the "Central Cerrado" route, following wetter areas in central Brazil. *N. clavipes* data do not support a northern connection through Caatinga, nor present exclusive lineages in NAF and SAF. Our findings indicate that, although common events might have affected the dynamics of rain forest-dwellers through the glacial cycles, the responses to these events are intrinsic of each species and strongly depends on environmental and ecological tolerances. Studying widely distributed species, which occur in more than one biome, has proven to be pivotal to a better understanding of diversification in highly complex regions.

ACKNOWLEDGEMENTS

The authors thank all colleagues who donated samples or helped with fieldwork: Arthur Costa, Bruno Rodrigues, Daniel Mendes, Douglas de Araujo, Estevam Silva, Fabián Salgado, Felipe Francisco, Karla Yotoko, Jair Mendes Jr., João Vasconcellos Neto, Luiz A. Silva, Maria Ligia Paseto, Milton Kanashiro, Paulo Motta, Raimundo Souto, Regiane Saturnino, Tércia Vargas, Thadeu Sobral-Souza, Thiago Kloss, Thierry Garnier. We would also like to thank Dr. Alexandre Bonaldo, in behalf of the Museu Paraense Emilio Goeldi, and Dr. Adalberto J. dos Santos, in behalf of the Coleção de Invertebrados em Meio Líquido da Universidade Federal de Minas Gerais, for donation of museum samples. We thank the Instituto Chico Mendes da Conservação da Biodiversidade (ICMBio) for collection permits (#27127 and 38889).

We also thank Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (fellowship #141040/2013-0 to LFMB) and São Paulo Research Foundation – FAPESP (grant #2012/02526-7 to VNS) for financial support.

REFERENCES

Auler, A.S., Wang, X., Edwards, R.L. *et al*. (2004) Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *Journal of Quaternary Science*, **19**, 693–701.

Bandelt, H.J., Forster P, Rohl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Bartoleti, L.F.M., Peres, E.A., Sobral-Souza, T. *et al*. (2017) Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal. *Journal of Biogeography*, in press. doi: 10.1111/jbi.12998

Batalha-Filho, H., Cabanne, G.S., Miyaki, C.Y. (2012) Phylogeography of an Atlantic forest passerine reveals demographic stability through the last glacial maximum. *Molecular Phylogenetics and Evolution*, **65**, 892–902.

Batalha-filho, H., Fjeldsa, J., Fabre, P., Miyaki, C.Y. (2013) Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of*

Ornithololology, **154**, 41–50.

Behling, H. (2002) South and southeast Brazilian grasslands during Late Quaternary times: a synthesis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **177**, 19–27.

Benham, P.M., Cuervo, A.M., McGuire, J.A., Witt, C.C. (2015) Biogeography of the Andean metaltail hummingbirds: contrasting evolutionary histories of tree line and habitat-generalist clades. *Journal of Biogeography*, **42**, 763-777.

Bidegaray-Batista, L., Arnedo, M.A. (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology*, **11**, 317.

Bouckaert, R., Heled, J. (2014) DensiTree 2: Seeing trees through the forest, doi: https://doi.org/10.1101/012401

Carnaval, A.C., Moritz, C. (2008). Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.

Carnaval, A.C., Waltari, E., Rodrigues, M.T. *et al*. (2014) Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20141461.

Cheng, H., Sinha, A., Cruz, F.W. *et al*. (2013) Climate change patterns in Amazonia and biodiversity. *Nature Communications*, **4**, 1411.

Christenhusz, M.J.M., Chase, M.W. (2013) Biogeographical patterns of plants in the Neotropics: dispersal rather than plate tectonics is most explanatory. *Botanical Journal of the Linnean Society*, **171**, 277–286.

Colgan, D.J., McLauchlan, A., Wilson, G.D.F. *et al*. (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology*, **46**, 419–437.

Cooper, M.A., Addison, F.T., Alvarez, R. *et al*. (1995) Basin development and tectonic history of the Llano basin, Eastern Cordillera, and middle Magdalena Valley, Colombia. *American Association of Petroleum Geologists Bulletin*, **79**, 1421–1443.

Corander, J., Sirén, J., Arjas, E. (2008) Bayesian Spatial Modelling of Genetic Population Structure. *Computational Statistics*, **23**, 111-129.

Costa, L.P. (2003) The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography*, **30**, 71–86.

Dantas, G.P.M., Sari, H.E.R., Cabanne, G.S. *et al*. (2015) Population genetic structure of the Atlantic Forest endemic Conopophaga lineata (Passeriformes: Conopophagidae) revelas a contact zone in the Atlantic Forest. *Journal of Ornithology*, **165**, 85-99.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9(8)**, 772.

De Ré, F.C., Gustani, E.C., Oliveira, A.P.F. *et al*. (2014) Brazilian populations of *Drosophila maculifrons* (Diptera: Drosophilidae): low diversity levels and signals of a population expansion after the Last Glacial Maximum. *Biological Journal of the Linnean Society*, **112**, 55–66.

D'Horta, F.M., Cabanne, G.S., Meyer, D., Miyaki, C.Y. (2011) The genetic effects of Late Quaternary climatic changes over a tropical latitudinal gradient: diversification of an Atlantic Forest passerine. *Molecular Ecology*, **20(9)**, 1923–1935.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.

Edgar, R.C. (2004) Muscle: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Excoffier, L., Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**(5), 294–299.

Fouquet, A., Cassini, C.S., Haddad, C.F.B., Pech, N., Rodrigues, M.T. (2014) Species delimitation, patterns of diversification and historical biogeography of the Neotropical frog genus *Adenomera*. *Journal of Biogeography*, **41**, 855–870.

Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gregory-Wodzicki, K.M. (2000) Uplift history of the central and northern Andes: a review. *Geological Society of America Bulletin*, **112**, 1091–1105.

Haffer, J. (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131-137.

Head, M.J., Gibbard, P.L. (2005) Early-Middle Pleistocene transitions: an overview and recommendation for the defining boundary. In: *Early-Middle Pleistocene transitions: the land-ocean evidence* (eds Head MJ and Gibbard PL), pp 1-18. The Geological Society of London.

Heled, J., Drummond, A.J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

Hoorn, C., Wesselingh, F.P., Steege, H. *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science*, **330**, 927–931.

Kumar, S., Stecher, G., Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870-1874.

Kuntner, M., Arnedo, M.A., Trontelj, P., Lokovsek, T., Agnarsson, I. (2013) A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution,* **69**, 961-979.

Lee, J.W., Jiang, L., Su, Y.C., Tso, I.M. (2004) Is central mountain range a geographic barrier to the giant wood spider *Nephila pilipes* (Araneae: Tetragnathidae) in Taiwan? A population genetic approach. *Zoological Studies*, **43**, 112-122.

Lee, V.M.J., Kuntner, M., Li, D. (2015) Ballooning behavior in the golden orbweb spider *Nephila pilipes* (Araneae:Nephilidae). *Frontiers in ecology and evolution*, **3**, 2:1-5.

Leigh, J.W., Bryant, D. (2015) popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

Leite, Y.L.R., Costa, L.P., Loss, A.C. *et al*. (2016) Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences*, **113**, 1008–1013.

Librado, P., Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Magalhaes, I.L., Oliveira, U., Santos, F.R. *et al*. (2014) Strong spatial structure, Pliocene diversification and cryptic diversity in the Neotropical dry forest spider *Sicarius cariri*. *Molecular Ecology*, **23**, 5323–5336.

Moore, C.W. (1977) The life cycle, habitat and variation in selected web parameters in the spider *Nephila clavipes* Koch (Araneidae). *American Midland Naturalist*, **98(1)**, 95–108.

Morley, R.J. (2000) Origin and evolution of tropical rain forests. John Wiley & Sons Ltd, Chichester.

Moritz, C., Richardson, K.S., Ferrier, S. *et al*. (2001) Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota.

Proceedings of the Royal Society of London-Series B: Biological Sciences, **268**, 1875-1881.

Muñoz-Ortiz, A., Velásquez-Álvarez, A.A., Guarnizo, C.E., Crawford, A. (2014) Of peaks and valleys: testing the roles of orogeny and habitat heterogeneity in driving allopatry in mid-elevation frogs (Aromobatidae: *Rheobates*) of the northern Andes. *Journal of Biogeography*, **42**, 193-205.

Oliveira-Filho, A.T., Ratter, J.A. (1995) A study of the origin of Central Brazilian forests by the analysis of plant species distribution patterns. *Edinburgh Journal of Botany*, **52**, 41–194.

Peres, E.A., Sobral-Souza, T., Perez, M.F. *et al.* (2015) Pleistocene niche stability and lineage diversification in the subtropical spider *Araneus omnicolor* (Araneidae). *PLoS ONE*, **10**, e0121543.

Peres, E.A., Silva, M.J., Solferini, V.N. (2017) Phylogeography of the spider *Araneus venatrix* (Araneidae) suggests past connections between Amazon and Atlantic rainforests. *Biological Journal of the Linnean Society*, doi: https://doi.org/10.1093/biolinnean/blx036

Por, F.D. (1992) Sooretama: the Atlantic rain forest of Brazil. The Hague: SPB Academic Publishing.

Prates, I., Rivera, D., Rodrigues, M.T., Carnaval, A.C. (2016). A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Journal of Biogeography*, **25**, 5174–5186.

Rambaut, A. (2012) *FigTree, version 1.4*, available at http://tree.bio.ed.ac.uk/software/figtree.

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J. (2014) *Tracer, version 1.6*, Available at http://beast.bio.ed.ac.uk/Tracer

Renner, S. (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, **165**, S4: S23-S33.

Rodrigues, M.T., Bertolotto, C.E.V., Amaro, R.C. *et al.* (2014) Molecular phylogeny, species limits, and biogeography of the Brazilian endemic lizard genus *Enyalius* (Squamata: Leiosauridae): an example of the historical relationship between Atlantic Forests and Amazonia. *Molecular Phylogenetics and Evolution*, **81**, 137–46.

Sobral-Souza, T., Lima-Ribeiro, M.S., Solferini, V.N. (2015). Biogeography of Neotropical rainforests: past connections between Amazon and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, doi: 10.1007/s10682-015-9780-9.

Su, Y.-C., Chang, Y.-H., Lee, S.-C., Tso, I.-M. (2007) Phylogeography of the giant wood spider (*Nephila pilipes*, Araneae) from Asian–Australian regions. *Journal of Biogeography*, **34,** 177–191.

Su, Y.-C., Chang, Y.-H., Smith, D. *et al*. (2011) Biogeography and speciation patterns of the golden orb spider genus *Nephila* (Araneae:Nephilidae) in Asia. *Zoological Science*, **28**(1), 47-55.

Tajima, F. (1989) Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Toews, D.P.L., Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Turchetto-Zolet, A.C., Pinheiro, F., Salgueiro, F., Palma-Silva, C. (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Valderrama, E., Perez-Eman, J.L., Brumfield, R.T., Cuervo, A.M., Cadena, C.D. (2014) The influence of the complex topography and dynamic history of the Andes on the evolutionary differentiation of a montane forest bird (*Premnoplex brunnescens*, Furnariidae). *Journal of Biogeography*, **41**, 1533–1546.

Vanzolini, P.E., Williams, E.E. (1981) The vanishing refuge: A mechanism for ecogeographic speciation. *Papéis avulsos de Zoologia*, **34(23)**, 251-255.

White, T.J., Bruns, T., Lee, S., Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, *et al*., eds. *PCR Protocols: A Guide to Methods and Applications*. New York: Academic Press, 315–322.

SUPPLEMENTARY INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Details on methodology.

Appendix S2 Supplementary tables.

Appendix S3 Supplementary figures.

DATA ACCESSIBILITY

- DNA sequences: GenBank, accession numbers XX0000-XX0000 (COI), XX0000-XX0000 (H3a), XX0000-XX0000 (ITS2).
- Geospatial coordinates of sampling locations, python scripts for coalescence simulations and xml files for Bayesian analyses: https://github.com/luizbartoleti/Nclavipes

BIOSKETCH

Luiz Filipe de Macedo Bartoleti is interested in unveiling the recent biogeography of the Neotropical region through the phylogeographical study of widely distributed species. All authors share a common interest in evolutionary biology and biogeography.

Author contributions: LFMB and VNS designed the research. LFMB, EAP and FVHMF carried out fieldwork. LFMB and MJS performed the research. MJS and VNS contributed with reagents and analytical tools. LFMB, EAP, FVHMF and VNS analyzed the data. LFMB, EAP, FVHMF and VNS wrote the paper. All authors have read and approved the final version of this manuscript.

TABLES

Table 1: Sampling locations of *Nephila clavipes* with details on biome and number of sequences obtained for each marker, per population.

Location	Biome	COI	H3a	ITS2
Alcobaça /BA (ALC)	NAF	8	10	10
Alta Floresta/MT (AFL)	AM	8	12	12
Aracruz/ES (ARC)	SAF	8	10	4
Belém/PA (BLM)	AM	8	-	-
Belo Horizonte/MG (BHR)	DD	8	16	14
Bragança/PA (BRG)	AM	8	-	-
Brasília/DF (BRS)	DD	8	14	14
Campo Grande/MS (CGD)	DD	1	-	-
Catalão/GO (CTL)	DD	1	-	-
Colombia – Amazon (CLA)	AM	1	2	2
Colombia – V. del Magdalena (CLM)	AM	6	12	12
Colombia – Pacific (CLP)	TAC	3	6	6
Colombia – V. del Cauca (CLC)	TAC	2	4	4
Florestópolis/PR (FLO)	SAF	8	-	-
Foz do Iguaçu/PR (FDI)	SAF	8	8	8
Goiânia/GO (GOI)	DD	3	-	-
Gramado/RS (GRA)	SAF	8	14	12
Ibirama/SC (IBM)	SAF	8	-	-
Ibitinga/SP (IBT)	DD	8	-	-
Ilha do Cardoso/SP (IDC)	SAF	8	-	-
Itaguaçu/ES (ITG)	SAF	3	-	-
Juiz de Fora/MG (JDF)	SAF	8	16	12
Jundiaí/SP (JDI)	SAF	8	-	-
Linhares/ES (LNH)	SAF	8	-	-
Macapá/AP (MAC)	AM	3	-	-
Manaus/AM (MAN)	AM	8	10	6
Mata de São João/BA (MSJ)	NAF	8	10	6
Moju/PA (MOJ)	AM	8	10	8
Monte Alegre do Sul/SP (MON)	SAF	8	-	-
Paraty/RJ (PTY)	SAF	7	10	2
Passa Quatro/MG (PQT)	SAF	8	-	-
Pinhalzinho/SP (PIN)	SAF	8	14	12
Pirenópolis/GO (PRN)	DD	8	10	8
Poços de Caldas/MG (PDC)	SAF	8	-	-
Porangatu/GO (PRG)	DD	8	10	2
Porto Velho/RO (PVL)	AM	6	-	-
Recife/PE (RCF)	NAF	3	-	-
Rio Branco/AC (RBC)	AM	8	10	8
Rio Claro/SP (RCL)	DD	8	-	-
Rio Preto do Eva/AM (RPE)	AM	1	-	-
Santa Teresa/ES (STE)	SAF	1	-	-
Santarém/PA (STM)	AM	8	-	-
Santos/SP (STO)	SAF	8	-	-
São Carlos/SP (SCA)	DD	8	-	-
Sapiranga/RS (SPG)	SAF	8	-	-
Silva Jardim/RS (SJD)	SAF	7	-	-
Teodoro Sampaio (TDS)	SAF	8	8	8
Uberlândia/MG (UBE)	DD	8	16	12
Viçosa/MG (VIC)	SAF	8	14	16
Total		320	246	198

Table 2: Genetic diversity indices for *Nephila clavipes*, calculated according to genetic groups and biomes. N = number of sequences; h = number of haplotypes; ss = number of segregating sites; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; D = Tajima's D; F_{ST} = fixation index, sd = standard deviation. * $p < 0.05$; ** $p < 0.02$.

COI								
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**	**F_{ST}**
Group I	140	27	30	0.7955 (0.0302)	0.0026 (0.0017)	-23.56**	-2.1045*	-
Group II	68	15	16	0.8292 (0.0296)	0.0028 (0.0018)	-7.464**	-1.4973*	-
Group III	58	9	11	0.7114 (0.0341)	0.0035 (0.0022)	-0.7814	-0.3577	-
Group IV	49	10	9	0.7185 (0.0575)	0.0023 (0.0016)	-3.8365	-0.9371	-
Group V	5	2	3	0.6000 (0.1753)	0.0030 (0.0024)	2.4290	1.5727	-
AM	73	19	30	0.8516 (0.0272)	0.0044 (0.0026)	-7.686**	-1.8148*	0.349
DD	69	23	25	0.9224 (0.0155)	0.0093 (0.0050)	-4.7852	0.2064	0.344
NAF	19	4	3	0.7076 (0.0739)	0.0014 (0.0012)	-0.3321	0.0967	0.571
SAF	154	31	40	0.8848 (0.0157)	0.0069 (0.0038)	-10.63**	-1.2593	0.540
TAC	5	2	3	0.6000 (0.1753)	0.0030 (0.0024)	2.4290	1.5727	0.500
Total	320	63	88	0.9375 (0.0068)	0.0111 (0.0058)	-24.50**	-1.5578*	0.751
H3a								
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**	**F_{ST}**
Group I	68	13	10	0.7204 (0.0480)	0.0092 (0.0057)	-3.1941	0.17883	-
Group II	48	11	7	0.7739 (0.0519)	0.0091 (0.0057)	-2.4119	1.06402	-
Group III	46	6	6	0.6300 (0.0633)	0.0081 (0.0052)	0.9730	1.14741	-
Group IV	74	9	8	0.6716 (0.0341)	0.0071 (0.0046)	-1.1094	0.14073	-
Group V	10	1	0	0	0	-	0	-
AM	56	12	7	0.8071 (0.0419)	0.0093 (0.0058)	-2.7395	1.26448	0.071
DD	66	8	6	0.6653 (0.0367)	0.0073 (0.0047)	-0.4626	0.97257	-0.038
NAF	20	4	5	0.7211 (0.0653)	0.0092 (0.0059)	2.13121	1.78515	0.080
SAF	94	17	12	0.6882 (0.0431)	0.0084 (0.0053)	-6.702**	-0.34514	0.134
TAC	10	1	0	0	0	-	0	0
Total	246	24	14	0.7428 (0.0224)	0.0089 (0.0055)	-10.01**	-0.14754	0.243
ITS2								
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**	**F_{ST}**
Group I	48	5	3	0.6144 (0.0524)	0.0039 (0.0031)	-0.6915	0.4445	-
Group II	40	5	4	0.609 (0.0662)	0.0034 (0.0029)	-1.2205	-0.568	-
Group III	40	5	3	0.5936 (0.0724)	0.0038 (0.0031)	-0.9054	0.3125	-
Group IV	60	5	4	0.5158 (0.0515)	0.0029 (0.0026)	-1.2442	-0.5985	-
Group V	10	1	0	0	0	-	0	-
AM	48	5	4	0.5895 (0.0592)	0.0032 (0.0028)	-1.1846	-0.5555	0.111
DD	50	6	4	0.591 (0.0579)	0.0035 (0.003)	-1.9206	-0.3628	-0.025
NAF	16	4	3	0.675 (0.0853)	0.0043 (0.0035)	-0.4728	0.0136	0.099
SAF	74	5	3	0.5424 (0.0494)	0.0034 (0.0029)	-0.6647	0.3368	-0.106
TAC	10	1	0	0	0	-	0	0
Total	198	7	14	0.6132 (0.0294)	0.0090 (0.0057)	1.8787	-0.5131	0.716

Table 3: Results of the Analysis of Molecular Variance (AMOVA) for *Nephila clavipes*, for each molecular marker. We performed two analysis: one considering the five biomes (TAC, AM, DD, SAF and NAF), in the left column; other excluding TAC individuals from the analysis. * $p < 0.05$.

	% of Variation	
Source of variation	**With TAC**	**W/o TAC**
COI		
Among biomes	49.15*	38.62*
Among populations within biomes	22.43*	27.09*
Within populations	28.42*	34.29*
H3a		
Among biomes	10.50*	-1.21
Among populations within biomes	10.19*	11.55*
Within populations	79.31*	89.65*
ITS2		
Among biomes	69.49*	-2.5
Among populations within biomes	2.37*	7.97*
Within populations	28.13*	94.53*

Table 4: Results of the model selection performed through ABC. We present the Posterior Probability (PP) for each model within scenarios (third column), except for Scenarios 1 and 2, each presenting only one model. The best model in each scenario (in bold) was used to perform the analysis Among Scenarios (fourth column). The model with highest PP in this analysis (in bold) was considered the best-fitting model for our data.

Scenario	Model	Within scenarios	Among scenarios
1	**1**	-	0.0000
2	**2**	-	0.0000
3	**3**	**0.6435**	0.0001
	4	0.0081	-
	5	0.3484	-
4	**6**	0.2275	-
	7	0.0072	-
	8	0.2538	-
	9	0.2146	-
	10	**0.2969**	**0.7493**
5	**11**	0.1876	-
	12	0.0019	-
	13	0.0317	-
	14	0.2239	-
	15	**0.5549**	0.2506

FIGURE LEGENDS

Figure 1: (a) Map with *Nephila clavipes* sampling locations in Brazil and Colombia. The two boxes in the right show details in Colombia [above – displaying the sampling sites according to each of the three Cordilleras: Western Cordillera (WC), Central Cordillera (CC) and Eastern Cordillera (EC)] and in the Brazilian Southeastern Region (below – a transition zone between SAF and DD). Light gray areas in the map represent humid forests (highlighting AM and AF), and the hatched area represents Cerrado. Pie charts represent the frequency of mitochondrial lineages in each location. Circle size is proportional to the number of individuals in each location. (b) Haplotype networks for each molecular marker: COI (left), H3a (middle) and ITS2 (right). In the COI network, colored polygons show the five mitochondrial lineages inferred by BAPS. The dashed line shows the distinction between Trans-Andean Colombia samples and the others. Circle size is proportional to the frequency of each haplotype, and dashes represent mutational events. Haplotypes in the networks are colored according to their occurrence in each biome.

Figure 2: Demographic scenarios simulated for the evolutionary history of *Nephila clavipes* in South America. Scenario I presumes panmixia among all populations (model I); Scenario II assumes an ancient split from TAC samples coincident with the final uplift of the Andean Central Cordillera (model 2); In Scenario III, besides TAC separations, we also tested a recent divergence between AM and a supposedly panmictic population comprising AF+DD, and tested for absence of gene flow (model 3), current (model 4) or LGM migration (model 5); for Scenario IV, we considered that all biomes have diverged, with a more recent split between AF and DD, and tested for absence of gene flow (model 6), current uniform migration (model 7), current non-uniform migration (model 8), uniform migration in the LGM (model 9), and non-uniform migration in the LGM (model 10); in Scenario 5, we tested the emergence of three

Brazilian lineages at the same time, and tested for absence (model 11), current uniform (model 12), current non-uniform (model 13), LGM uniform (model 14) and LGM non-uniform (model 15) migration.

Figure 3: Bayesian phylogenetic inferences for *Nephila clavipes* individuals. In the box, the multilocus species tree showing the level of uncertainty among the three markers, mainly in the more recent Brazilian branches. In red, the most frequently recovered consensus tree. In the COI tree, we display the five mitochondrial lineages inferred by BAPS, with posterior probabilities for the main nodes. In the Group I, NAF individuals are highlighted with black lines in front of the tips. The circles show results from the biogeographical analysis of ancestral area reconstruction, representing the most likely biome of origin from each mitochondrial group.

Figure 4: Results for the Extended Bayesian Skyline Plot analyses for *Nephila clavipes*. Analyses were performed jointly for the three markers. Continuous lines represent the medians, and the gray shadings display 95% HPD for each analysis.

Figure 1

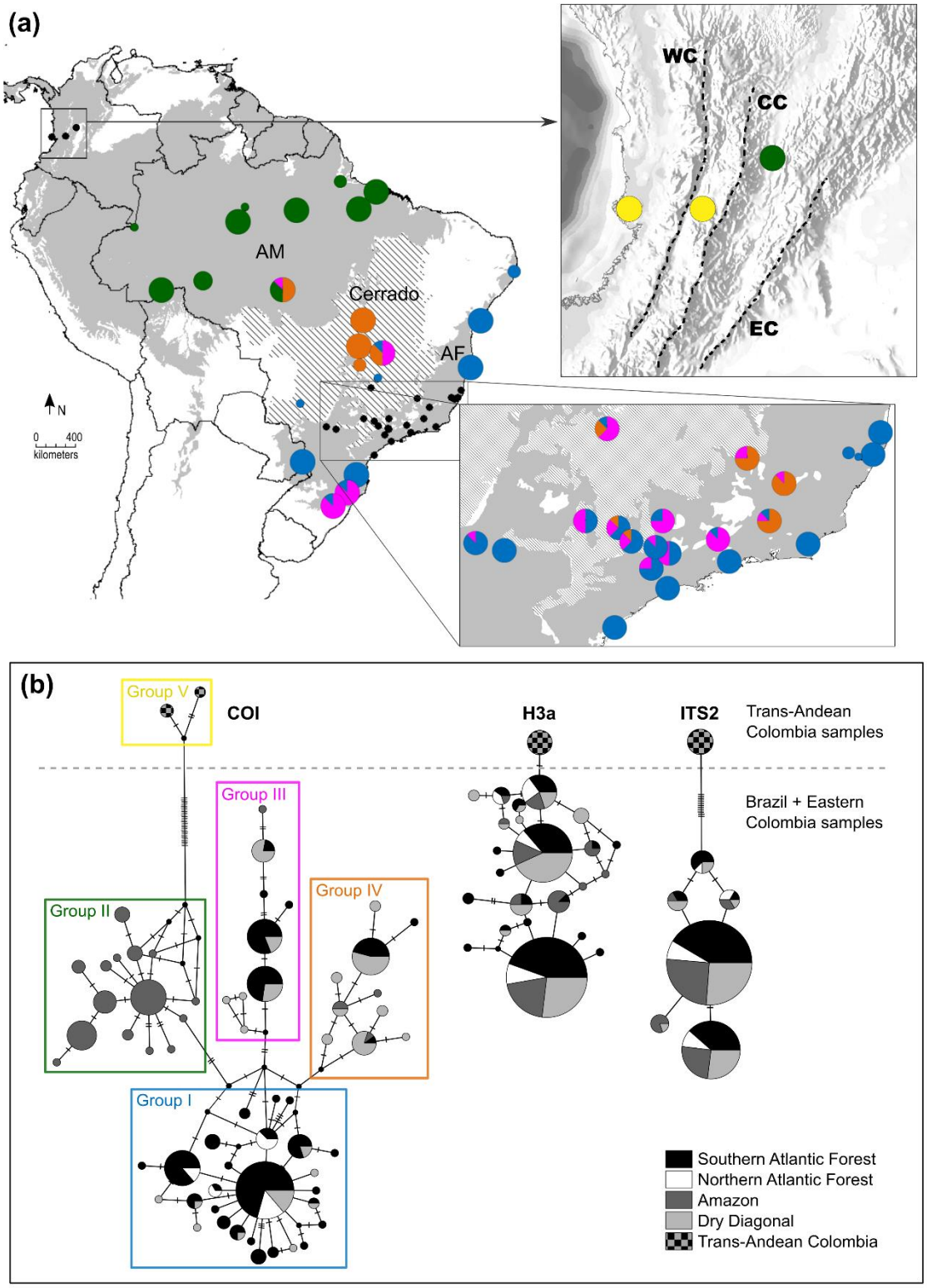

(a)
WC
CC
EC
AM
Cerrado
AF
N
0 400
kilometers
(b)
Group V
COI
H3a
ITS2
Trans-Andean
Colombia samples
Brazil + Eastern
Colombia samples
Group III
Group II
Group IV
Group I
Southern Atlantic Forest
Northern Atlantic Forest
Amazon
Dry Diagonal
Trans-Andean Colombia

Figure 2

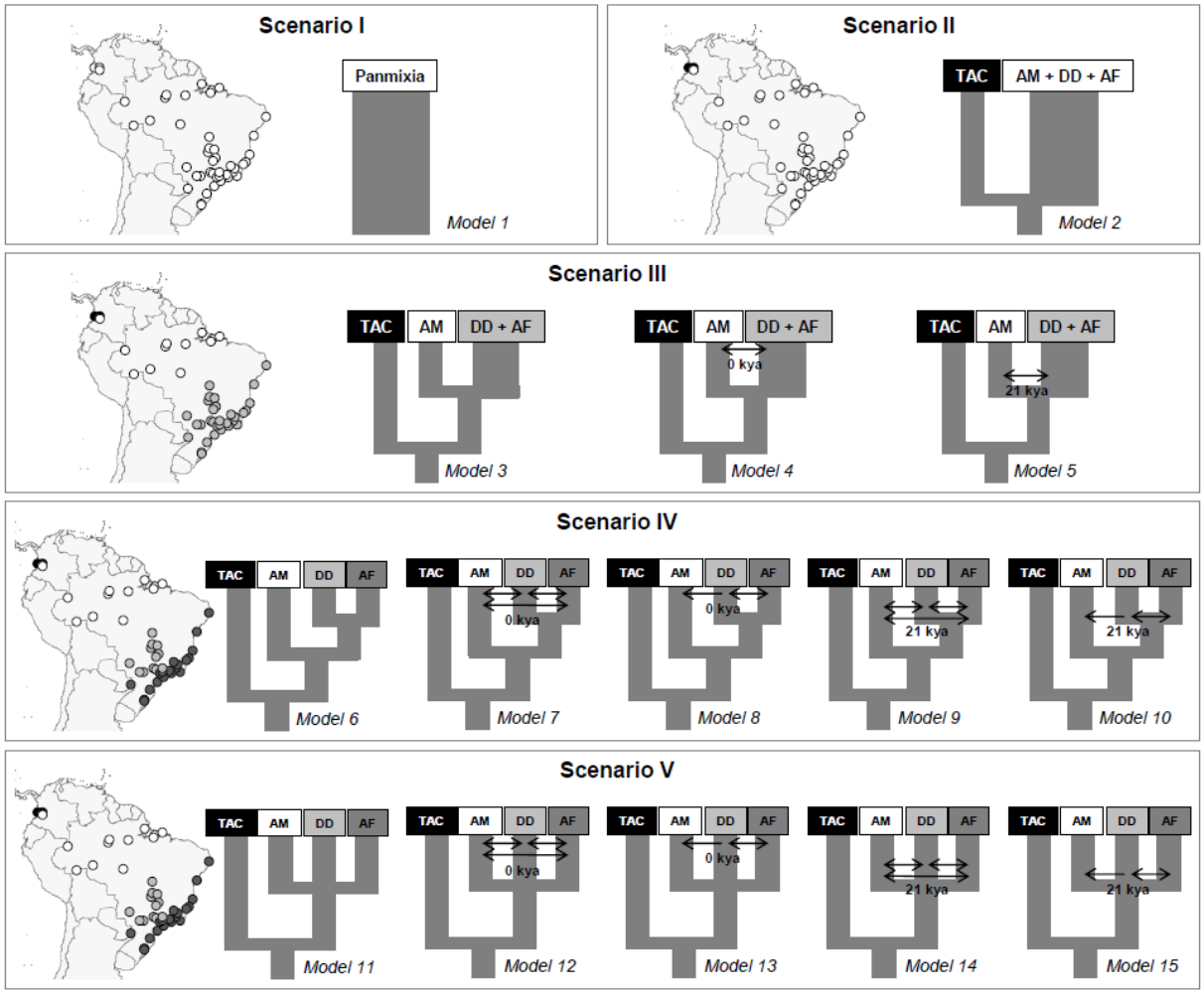

Scenario I
Panmixia
Model 1
Scenario II
TAC
AM + DD + AF
Model 2
Scenario III
TAC
AM
DD + AF
Model 3
0 kya
Model 4
21 kya
Model 5
Scenario IV
TAC
AM
DD
AF
Model 6
0 kya
Model 7
0 kya
Model 8
21 kya
Model 9
21 kya
Model 10
Scenario V
TAC
AM
DD
AF
Model 11
0 kya
Model 12
0 kya
Model 13
21 kya
Model 14
21 kya
Model 15

Figure 3

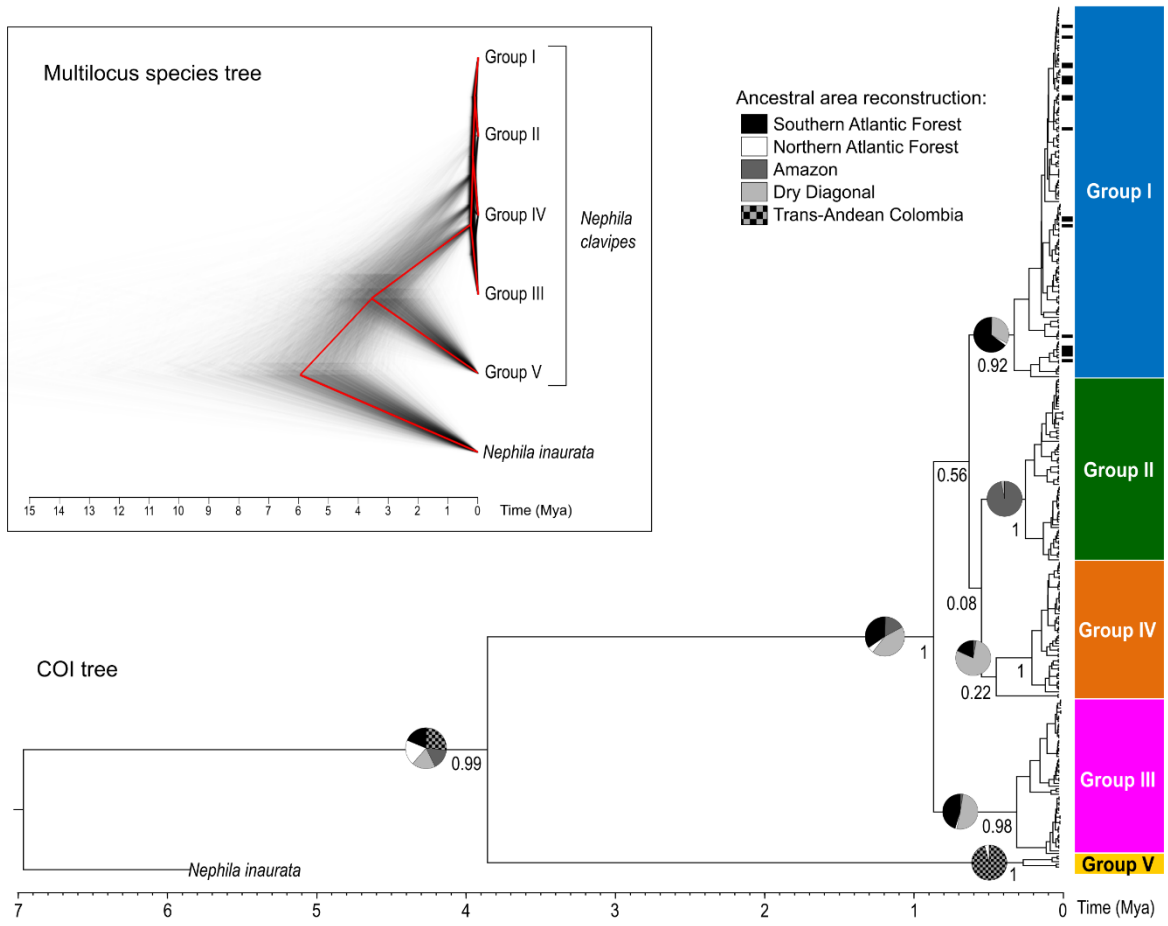
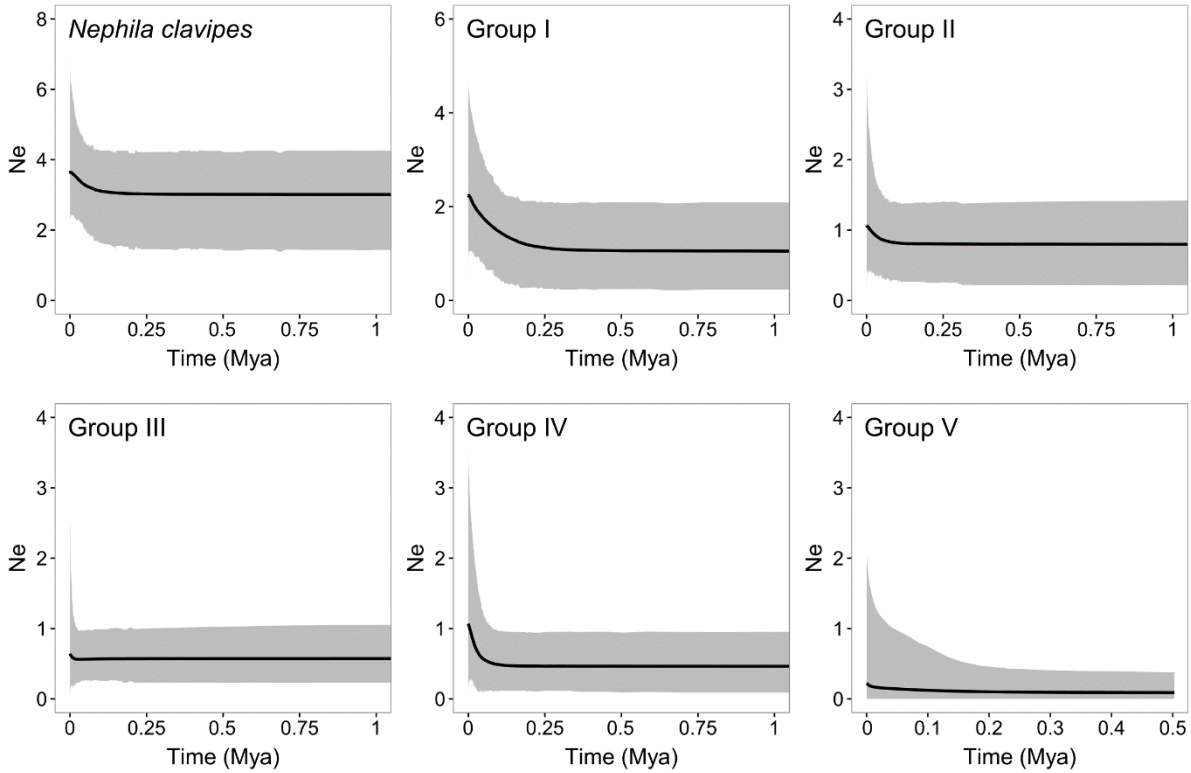

Multilocus species tree
Group I
Group II
Group IV
Group III
Group V
Nephila clavipes
Nephila inaurata
15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0 Time (Mya)
Ancestral area reconstruction:
Southern Atlantic Forest
Northern Atlantic Forest
Amazon
Dry Diagonal
Trans-Andean Colombia
COI tree
0.99
1
0.92
0.56
1
0.08
1
0.22
0.98
1
Nephila inaurata
Group I
Group II
Group IV
Group III
Group V
7 6 5 4 3 2 1 0 Time (Mya)

Figure 4


Nephila clavipes
Group I
Group II
Group III
Group IV
Group V
Ne
Time (Mya)

Journal of Biogeography

SUPPORTING INFORMATION

Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest

Luiz Filipe de M. Bartoleti, Elen Arroyo Peres, Fernanda von Hertwig Mascarenhas Fontes, Márcio José da Silva and Vera Nisaka Solferini

APPENDIX S1. DETAILS ON METHODOLOGY.

Model Testing

We used an Approximate Bayesian Computation (ABC; Beaumont, 2010; Csilléry *et al*, 2010; Sunnaker *et al*, 2013) framework to perform the model testing. We separated the individuals in the previously defined groups: TAC, AM, DD and AF (we kept NAF and SAF as one single population in this analysis due to genetic results) according to their geographical location. We constructed demographic models that represent biogeographical hypothesis concerning the evolutionary history of *N. clavipes* and the Neotropical region (Fig. 2). We grouped models in scenarios in order to first compare similar models (which represent alternatives for the same hypothesis) and then compare the best model of each scenario, to select which hypothesis was the best overall. We performed the analysis jointly for mitochondrial and nuclear markers.

The first scenario presents only one model, representing a single panmictic population with no demographic changes over time (Model 1). In this hypothesis, all sampled populations are connected and there are no barriers to gene flow in the sampled area, therefore neither geological nor climate events would have promoted diversification (TAC+AM+DD+AF).

The second scenario also presents only one model (Model 2); in this case, divergence in *N. clavipes* would have happened due to Neogene orogenic activity, mainly the Andean uplift. Therefore, the divergence between TAC individuals and the other samples would be coincident with the estimated time for the Central Cordillera final uplift (~10-6 Ma). The other individuals would constitute a single panmictic population (TAC(AM+DD+AF)).

For scenarios 3-5 we considered that, besides the divergence due to the Andean Uplift in the Colombian samples, climate fluctuations promoted diversification among Brazilian biomes during Quaternary. In each of these scenarios, we tested three mutually exclusive

hypotheses: i) populations experienced no migration among Brazilian biomes after the diversification; ii) despite moments of isolation, Brazilian populations are connected in the present through the net of gallery forests in Cerrado, represented by multidirectional migration; iii) habitats shifts in the LGM would have promoted connections among biomes; therefore, migration among populations would have occurred mainly during this period due to biome expansion/retraction.

In Scenario 3, Brazilian samples would present a split separating AM from DD and AF, the last two still constituting a panmictic population (TAC(AM(DD+AF))). We included three models in this scenario: Model 3 presents no gene flow among any of the populations; in Model 4 we considered present-day migration between AM and DD+AF; in Model 5, we considered migration during LGM.

In Scenario 4, AF and DD do not constitute a panmictic population, and diverged more recently than AM (TAC(AM(DD(AF)))). Model 6 represents absence of gene flow; Model 7 includes multidirectional migration among AM, AF and DD in the present; in Model 8 we considered current non uniform migration among these biomes, as suggested by our genetic results (see Results): DD would serve as source of migrants for both AM and AF, and AF would send migrants only to DD; in Model 9, we considered multidirectional migration among Brazilian biomes dating from the LGM; and in Model 10, the connections among these biomes also would have occurred in the LGM but they would not have been uniform, similar to model 8.

In Scenario 5, we tested a fragmentation from a past widespread Brazilian lineage. Therefore, Brazilian populations (AM, DD and AF) would present divergence in the Pleistocene, with three distinct lineages emerging at the same time (TAC(AM/DD/AF)). The five models within this scenario are similar to those from scenario 4: no gene flow (Model 11); uniform multidirectional migration in the present for Brazilian biomes (Model 12); uneven migration in the present where DD is the main source of migrants (Model 13); multidirectional migration in the LGM (Model 14); uneven migration in the LGM (Model 15).

We estimated N_e for the species through the formula $N_e = \theta/\mu$ using the mitochondrial data (θ was estimated on DnaSP) with a substitution rate of 0.0112 substitutions/site/million years (Kuntner *et al*, 2013), as this is the more reliable substitution rate for any nephilid gene. This information was used to calculate the time of past events in units of $4N_0$ generations through the formula τ = Number of generations/$4*N_e$, considering the generational time as one year.

We performed 100,000 simulations per model on *ms* (Hudson, 2002) using custom Python scripts. We set parameters (θ and Migration) as flat uniform prior distributions to allow a wider prior sampling as an exploratory step. Time of divergence between TAC and the Brazilian samples (in scenarios 2-5) was given a uniform distribution that represented the Central Cordillera final uplift (10-6 Ma). Divergences between Brazilian samples in scenarios 3-5 were given a uniform distribution which represented the Pleistocene period (2.588-0.011 Ma). We then calculated the summary statistics [total nucleotide diversity, number of segregating sites, Tajima's D, nucleotide diversity within and between populations (π, ss, D, π_w, π_b, respectively)] of the newly simulated data using a PERL script by N. Takebayashi (available at http://raven.iab.alaska.edu/~ntakebay/teaching/programming/coalsim/scripts/msSS.pl), grouped them in vectors and found the one that optimizes model selection through a rejection step performed with pseudo-observed data.

We narrowed the parameters distribution using the summary statistics selected and the "abc" package (Csilléry *et al*, 2012) in R, using the 0.1% simulations that best fit the empirical data. We then used the posterior distribution obtained in these exploratory runs as priors for 400,000 new simulations for each model. We performed model selection in R using the "abc" package, keeping the 0.01% of the sequences more similar to our empirical data based on a neural network rejection. We performed this procedure within scenarios 3-5, selected the best model within each scenario, then gathered these models with scenarios 1 and 2 and contrasted the five models to find the one with higher posterior probability to fit our empirical data.

References:

Beaumont MA (2010) Approximate Bayesian Computation in Evolution and Ecology. *Annual Reviews*, **41**, 379-406.

Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, **25**, 410-418.

Csilléry K, Francois O, Blum MGB (2012) abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution*, **3**, 475-479.

Hudson, RR (2002) Generating samples under a Wright-Fisher neutral model. *Bioinformatics*, **18**, 337-338.

Kuntner M, Arnedo MA, Trontelj P, Lokovsek T, Agnarsson I (2013) A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution*, **69**, 961-979.

Sunnaker M, Busetto AG, Numminem E *et al*. (2013) Approximate Bayesian Computation. PLoS Computational Biology, doi: http://dx.doi.org/10.1371/journal.pcbi.1002803.

Journal of Biogeography

SUPPORTING INFORMATION

Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest

Luiz Filipe de M. Bartoleti, Elen Arroyo Peres, Fernanda von Hertwig Mascarenhas Fontes, Márcio José da Silva and Vera Nisaka Solferini

APPENDIX S2. SUPPLEMENTARY TABLES.

Table S1: COI genetic diversity indices for *Nephila clavipes* populations. N = number of sequences; h = number of haplotypes; ss = number of segregating sites; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; D = Tajima's D; F_{ST} = fixation index, sd = standard deviation. * $p < 0.05$; ** $p < 0.02$.

COI							
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**
Alcobaça /BA (ALC)	8	4	3	0.821 (0.101)	0.0023 (0.0018)	-0.569	0.839
Alta Floresta/MT (AFL)	8	6	18	0.928 (0.084)	0.0124 (0.0074)	0.315	0.336
Aracruz/ES (ARC)	8	1	0	0	0	-	0
Belém/PA (BLM)	8	6	5	0.893 (0.111)	0.0024 (0.0018)	-3.589*	-1.175
Belo Horizonte/MG (BHR)	8	4	9	0.75 (0.139)	0.0067 (0.0042)	1.631	0.696
Bragança/PA (BRG)	8	2	1	0.25 (0.18)	0.0004 (0.0005)	-0.182	-1.055
Brasília/DF (BRS)	8	6	17	0.893 (0.111)	0.0115 (0.0069)	0.163	0.236
Campo Grande/MS (CGD)	1	1	0	1	0	-	0
Catalão/GO (CTL)	1	1	0	1	0	-	0
Colombia – Amazon (CLA)	1	1	0	1	0	-	0
Colombia – V. del Magdalena (CLM)	6	1	0	0	0	-	0
Colombia – Pacific (CLP)	3	2	3	0.667 (0.314)	0.0034 (0.0032)	1.609	0
Colombia – V. del Cauca (CLC)	2	1	0	0	0	-	0
Florestópolis/PR (FLO)	8	3	5	0.679 (0.122)	0.0026 (0.0020)	1.018	-0.923
Foz do Iguaçu/PR (FDI)	8	2	1	0.25 (0.18)	0.0004 (0.0005)	-0.182	-1.055
Goiânia/GO (GOI)	3	2	2	0.667 (0.314)	0.0022 (0.0023)	1.061	0
Gramado/RS (GRA)	8	3	8	0.464 (0.2)	0.0034 (0.0024)	1.572	-1.701*
Ibirama/SC (IBM)	8	1	0	0	0	-	0
Ibitinga/SP (IBT)	8	4	7	0.786 (0.113)	0.0056 (0.0037)	1.25	1.158
Ilha do Cardoso/SP (IDC)	8	4	4	0.643 (0.184)	0.0017 (0.0014)	-1.236	-1.535*
Itaguaçu/ES (ITG)	3	1	0	0	0	-	0
Juiz de Fora/MG (JDF)	8	4	14	0.643 (0.184)	0.0074 (0.0046)	1.877	-0.949
Jundiaí/SP (JDI)	8	5	9	0.857 (0.108)	0.0059 (0.0038)	0.059	0.091
Linhares/ES (LNH)	8	3	3	0.679 (0.122)	0.0022 (0.0017)	0.723	0.585
Macapá/AP (MAC)	3	1	0	0	0	-	0
Manaus/AM (MAN)	8	2	2	0.429 (0.169)	0.0014 (0.0013)	1.653	0.414
Mata de São João/BA (MSJ)	8	1	0	0	0	-	0
Moju/PA (MOJ)	8	7	8	0.964 (0.077)	0.0042 (0.0028)	-3.802*	-0.917
Monte Alegre do Sul/SP (MON)	8	4	7	0.643 (0.184)	0.0032 (0.0023)	0.081	-1.359
Paraty/RJ (PTY)	7	3	4	0.667 (0.16)	0.0037 (0.0026)	1.508	1.633
Passa Quatro/MG (PQT)	8	2	6	0.25 (0.18)	0.0025 (0.0019)	2.822	-1.64*
Pinhalzinho/SP (PIN)	8	5	9	0.893 (0.086)	0.0071 (0.0045)	0.427	1.049
Pirenópolis/GO (PRN)	8	5	5	0.893 (0.086)	0.0032 (0.0023)	-1.232	0.0005
Poços de Caldas/MG (PDC)	8	4	7	0.75 (0.14)	0.0047 (0.0032)	0.869	0.214
Porangatu/GO (PRG)	8	3	2	0.464 (0.2)	0.0011 (0.0011)	-0.478	-0.448
Porto Velho/RO (PVL)	6	3	4	0.733 (0.155)	0.0031 (0.0024)	0.892	0.355
Recife/PE (RCF)	3	1	0	0	0	-	0
Rio Branco/AC (RBC)	8	6	6	0.929 (0.084)	0.0028 (0.0021)	-3.114*	-1.28
Rio Claro/SP (RCL)	8	7	12	0.964 (0.077)	0.0074 (0.0046)	-2.221	-0.256
Rio Preto do Eva/AM (RPE)	1	1	0	1	0	-	0
Santa Teresa/ES (STE)	1	1	0	1	0	-	0
Santarém/PA (STM)	8	3	2	0.464 (0.2)	0.0008 (0.0009)	-0.999	-1.310
Santos/SP (STO)	8	5	4	0.857 (0.108)	0.0022 (0.0017)	-2.169*	-0.727
São Carlos/SP (SCA)	8	5	12	0.857 (0.108)	0.0069 (0.0044)	0.373	-0.566
Sapiranga/RS (SPG)	8	3	8	0.464 (0.2)	0.0034 (0.0024)	1.573	-1.701*
Silva Jardim/RS (SJD)	7	4	10	0.714 (0.181)	0.0062 (0.0041)	1.092	-0.483
Teodoro Sampaio (TDS)	8	3	7	0.464 (0.2)	0.0029 (0.0022)	1.286	-1.674*
Uberlândia/MG (UBE)	8	4	14	0.821 (0.101)	0.01 (0.0061)	2.645	0.499
Viçosa/MG (VIC)	8	3	11	0.464 (0.2)	0.0052 (0.0034)	2.631	-1.336

Table S2: H3a and ITS2 genetic diversity indices for *Nephila clavipes* populations. N = number of sequences; h = number of haplotypes; ss = number of segregating sites; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; D = Tajima's D; F_{ST} = fixation index, sd = standard deviation. * $p < 0.05$; ** $p < 0.02$.

H3a							
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**
Alcobaça /BA (ALC)	10	3	4	0.622 (0.138)	0.008 (0.0086)	1.927	1.471
Alta Floresta/MT (AFL)	12	6	5	0.864 (0.072)	0.008 (0.0055)	-1.412	0.632
Aracruz/ES (ARC)	10	5	7	0.844 (0.08)	0.0122 (0.0079)	0.305	0.862
Belo Horizonte/MG (BHR)	16	4	4	0.642 (0.081)	0.0068 (0.0047)	0.914	1.154
Brasília/DF (BRS)	14	5	5	0.67 (0.126)	0.0091 (0.006)	0.323	1.402
Colombia - Amazon (CLA)	2	2	3	1 (0.5)	0.0123 (0.0142)	1.099	0
Colombia - V. del Magdalena (CLM)	12	3	3	0.318 (0.164)	0.0032 (0.0028)	0.18	-0.729
Colombia - Pacific (CLP)	6	1	0	0	0	-	0
Colombia - V. del Cauca (CLC)	4	1	0	0	0	-	0
Foz do Iguaçu/PR (FDI)	8	2	3	0.25 (0.18)	0.0031 (0.0029)	1.415	-1.447
Gramado/RS (GRA)	14	4	5	0.692 (0.094)	0.0083 (0.0056)	1.221	0.99
Juiz de Fora/MG (JDF)	16	6	5	0.6167 (0.135)	0.0046 (0.0035)	-2.337	-0.853
Manaus/AM (MAN)	10	4	6	0.8 (0.089)	0.011 (0.0072)	1.257	1.064
Mata de São João/BA (MSJ)	10	4	5	0.822 (0.072)	0.0095 (0.0064)	0.92	1.233
Moju/PA (MOJ)	10	6	5	0.889 (0.075)	0.0079 (0.0055)	-1.953	0.326
Paraty/RJ (PTY)	10	6	6	0.778 (0.137)	0.0086 (0.0059)	-1.718	-0.062
Pinhalzinho/SP (PIN)	14	3	6	0.483 (0.142)	0.0094 (0.0061)	2.999	0.759
Pirenópolis/GO (PRN)	10	4	4	0.733 (0.12)	0.0073 (0.0052)	0.334	0.988
Porangatu/GO (PRG)	10	4	4	0.733 (0.1)	0.0075 (0.0053)	0.388	1.109
Rio Branco/AC (RBC)	10	4	5	0.644 (0.152)	0.0092 (0.0063)	0.867	1.082
Teodoro Sampaio (TDS)	8	3	3	0.607 (0.164)	0.0057 (0.0045)	0.825	0.839
Uberlândia/MG (UBE)	16	3	4	0.575 (0.08)	0.0069 (0.0048)	2.286	1.218
Viçosa/MG (VIC)	14	3	4	0.472 (0.136)	0.006 (0.0043)	1.728	0.533
ITS2							
Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**
Alcobaça /BA (ALC)	10	4	3	0.733 (0.12)	0.0048 (0.004)	-0.83	-0.13
Alta Floresta/MT (AFL)	12	2	1	0.53 (0.076)	0.0025 (0.0025)	1.152	1.381
Aracruz/ES (ARC)	4	3	3	0.833 (0.222)	0.0079 (0.0069)	-0.133	0.168
Belo Horizonte/MG (BHR)	14	4	3	0.659 (0.09)	0.0043 (0.0035)	-0.647	-0.137
Brasília/DF (BRS)	14	4	3	0.626 (0.11)	0.004 (0.0034)	-0.784	-0.315
Colombia - Amazon (CLA)	2	2	1	1 (0.5)	0.0047 (0.0067)	0	0
Colombia - V. del Magdalena (CLM)	12	2	1	0.303 (0.147)	0.0014 (0.0018)	0.297	-0.194
Colombia - Pacific (CLP)	6	1	0	0	0	-	0
Colombia - V. del Cauca (CLC)	4	1	0	0	0	-	0
Foz do Iguaçu/PR (FDI)	8	3	2	0.464 (0.2)	0.0024 (0.0025)	-0.999	-1.31
Gramado/RS (GRA)	12	4	3	0.758 (0.081)	0.0062 (0.0047)	-0.05	1.022
Juiz de Fora/MG (JDF)	12	2	1	0.409 (0.133)	0.0019 (0.0021)	0.735	0.54
Manaus/AM (MAN)	6	3	2	0.733 (0.155)	0.0044 (0.0040)	-0.304	0.311
Mata de São João/BA (MSJ)	6	2	1	0.6 (0.129)	0.0028 (0.003)	0.795	1.445
Moju/PA (MOJ)	8	2	1	0.571 (0.094)	0.0027 (0.0028)	0.966	1.444
Paraty/RJ (PTY)	2	2	3	1 (0.5)	0.0142 (0.0164)	1.099	0
Pinhalzinho/SP (PIN)	12	2	1	0.485 (0.106)	0.0023 (0.0024)	1.003	1066
Pirenópolis/GO (PRN)	8	2	1	0.429 (0.169)	0.002 (0.0023)	0.536	0.333
Porangatu/GO (PRG)	2	2	2	1 (0.5)	0.0095 (0.0116)	0.693	0
Rio Branco/AC (RBC)	8	5	1	0.571 (0.094)	0.0027 (0.0028)	0.966	1.444
Teodoro Sampaio (TDS)	8	3	2	0.464 (0.2)	0.0024 (0.0025)	-0.999	-1.31
Uberlândia/MG (UBE)	12	3	2	0.53 (0.136)	0.0027 (0.0027)	-0.362	-0.382
Viçosa/MG (VIC)	16	2	1	0.458 (0.095)	0.0022 (0.0022)	1.096	1.034

Table S3-S5: Tables displaying pairwise FST values among populations for COI (S3), H3a (S4) and ITS2 (S5). Due to the large number of lines and columns, these tables are available as a .xlsx file on https://github.com/luizbartoleti/Nclavipes.

Table S6: Results from the ancestral area reconstruction for *Nephila clavipes*. Lines display the mitochondrial lineages and the species' root, and columns display the five discrete biomes considered. The most likely biome for each lineage origin is highlighted in bold.

	TAC	AM	DD	NAF	SAF
Group I	0.0006	0.0066	0.3223	0.0222	**0.6483**
Group II	0.0001	**0.9669**	0.0267	0.0012	0.0051
Group III	0.0008	0.0255	**0.5045**	0.0159	0.4533
Group IV	0.0000	0.0278	**0.7881**	0.0031	0.1810
Group V	**0.9554**	0.0017	0.0007	0.0414	0.0008
Brazilian samples	0.0055	0.1659	0.4335	0.0488	0.3463
N. clavipes	0.2648	0.1628	0.1865	0.1990	0.1869

Journal of Biogeography

SUPPORTING INFORMATION

Phylogeography of the widespread spider *Nephila clavipes* (Araneae: Araneidae) in South America reveals geological and climate diversification, and Pleistocene connections between Amazon and Atlantic Forest

Luiz Filipe de M. Bartoleti, Elen Arroyo Peres, Fernanda von Hertwig Mascarenhas Fontes, Márcio José da Silva and Vera Nisaka Solferini

APPENDIX S3. SUPPLEMENTARY FIGURES.

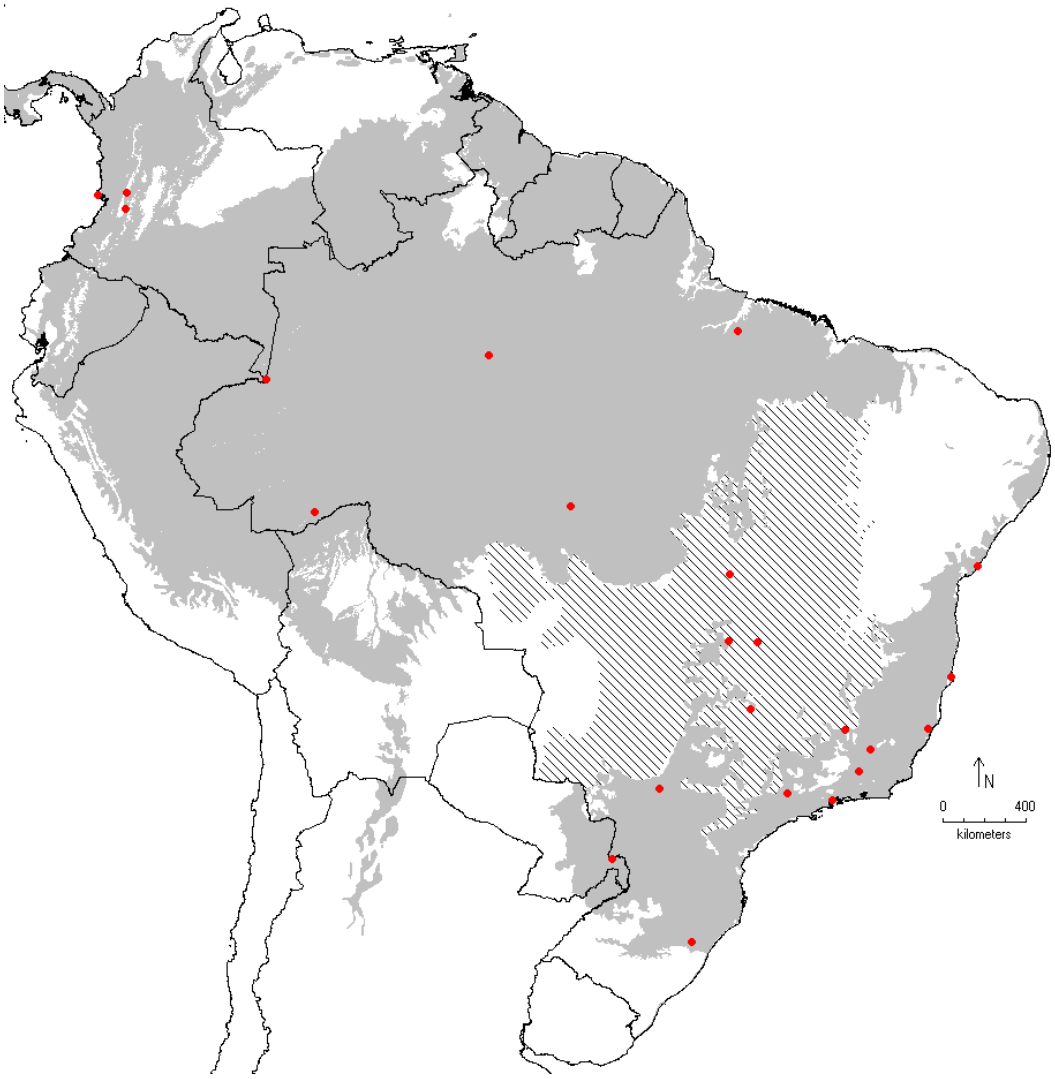


Figure S1: Map showing the locations that constituted the nuclear subsample.

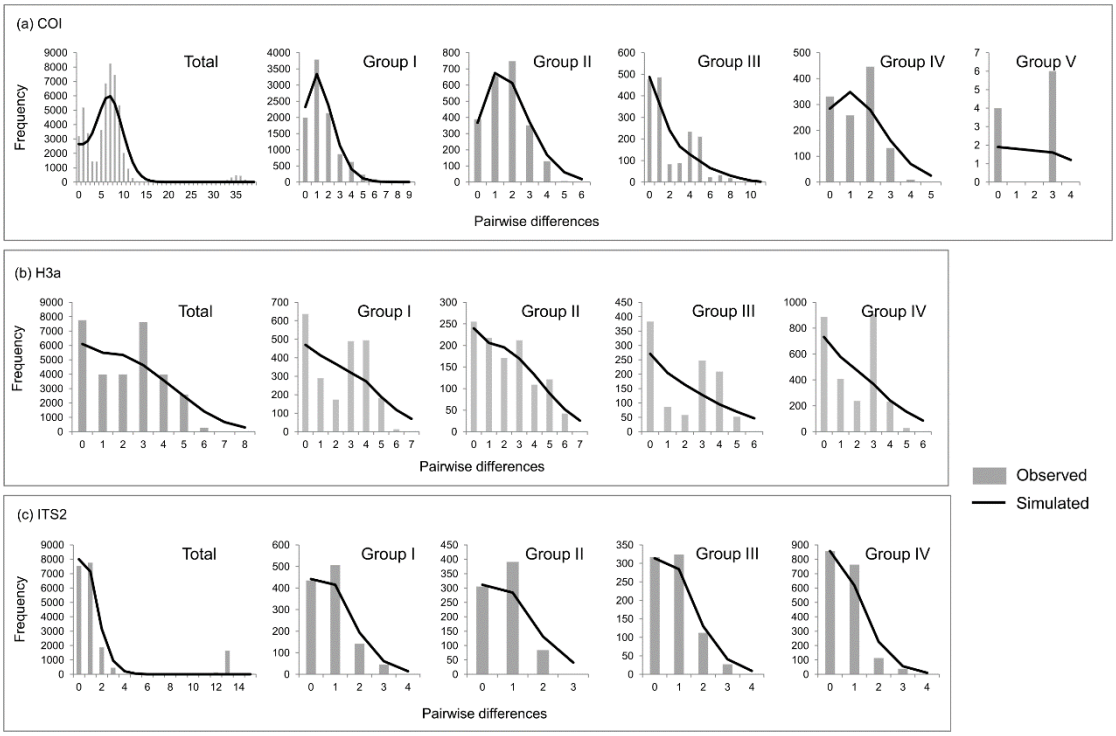


Figure S2: Results from the mismatch distribution analysis for the three markers separately – (a) COI, (b) H3a and (c) ITS2. The continuous lines represent the expected number of pairwise differences in a scenario of sudden expansion, and columns represent the actual number observed in our data. The analysis was performed in the overall dataset and for each mitochondrial group. Mismatch distributions were not calculated for Group V for the nuclear markers, as they presented only one haplotype.

MANUSCRITO II

Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal

(Manuscrito publicado no periódico *Journal of Biogeography*)

ORIGINAL ARTICLE:

Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal

Luiz Filipe de Macedo Bartoleti[1]; Elen Arroyo Peres[2]; Thadeu Sobral-Souza[3]; Fernanda von Hertwig Mascarenhas Fontes[1]; Márcio José da Silva[4]; Vera Nisaka Solferini[1]*

[1]Department of Genetics, Evolution and Bioagents, Institute of Biology, University of Campinas, Campinas, São Paulo, Brazil

[2]Department of Zoology, Institute of Biosciences, University of São Paulo, São Paulo, São Paulo, Brazil

[3]Spatial Ecology and Conservation Lab (LEEC), São Paulo State University, Rio Claro, São Paulo, Brazil

[4]Center of Molecular Biology and Genetic Engineering, University of Campinas, Campinas, São Paulo, Brazil

*Correspondence: Vera N. Solferini, Laboratory of Genetic Diversity, Department of Genetics, Evolution and Bioagents, Institute of Biology, University of Campinas, Rua Monteiro Lobato 255, 13083-862, Campinas, São Paulo, Brazil.

E-mail: veras@unicamp.br

Running header: Phylogeography of *Nephila sexpunctata*

Word count: 6534

Word count (abstract): 295

Number of pages for figures: 3-4.

ABSTRACT:

Aim The Neotropical Dry Diagonal (DD) is a corridor of distinct dry environments in central South America. The main hypotheses suggest that these environments may have expanded during glacial cycles together with a retraction of rainforests, and then shrank when the climate became wetter and warmer. However, few studies have explicitly tested hypotheses on DD evolution. We conducted a phylogeographical study on *Nephila sexpunctata*, a spider endemic to Neotropical dry habitats, to enrich the understanding of DD evolutionary history.

Location Southwestern portion of the DD (Cerrado and Chaco).

Methods We sequenced two mitochondrial regions and one nuclear DNA region of 65 individuals from 10 locations across the species' range. We conducted analyses of genetic structure, variability, demography and inferred divergence times. We used an ecological niche modelling framework to generate hypotheses on the species' distribution along the last glacial cycle and a model-based approach to test demographic scenarios that might explain the genetic patterns.

Results Both markers presented low genetic diversity. Mitochondrial markers had high genetic structure, with specific geographical lineages within each biome, while the nuclear marker presented low genetic structure. Phylogeographical and demographic events encompass the last glacial cycle, with a strong post-Last Glacial Maximum (LGM) population and spatial expansion. Model testing corroborated the recent demographic expansion of *N. sexpunctata*.

Main Conclusions Our data suggest that Quaternary climate cycles, when moist forests periodically expanded towards the dry vegetation, influenced diversification of DD organisms by promoting demographic events. These events might explain the genetic structure observed in *N. sexpunctata*; differently from what was expected by the Pleistocene refugia model, some DD species may have experienced post-LGM expansion. Climate and geologic events have both affected DD-endemic organisms' diversification.

Keywords Cerrado, Chaco, dry diagonal, ecological niche modelling, Last Glacial Maximum, Neotropics, *Nephila*, phylogeography, Pleistocene refugia model

INTRODUCTION

The Neotropics are the most diverse biogeographical region of the planet, presenting a great variety of ecoregions, from tropical rain forests to drier open formations. The largest continuous portion of dry vegetation in this region forms a northeast-southwest corridor that crosses South America, called the Dry Diagonal, composed of three biomes: Caatinga, the largest Neotropical nucleus of Seasonally Dry Tropical Forests (SDTFs) in northeastern Brazil; Cerrado, a mosaic of phytophysiognomies predominantly covered by savannas in the central part of the continent; and Chaco, composed mainly of subtropical dry forests and open woodlands in northern Argentina, Paraguay and Bolivia (Fig. 1). Evidence suggests that in the Eocene, when the climate was wetter and warmer than today, this area was covered by wet forests; the Amazon and Atlantic Forest were connected, forming a rain forest belt through most of South America (Morley, 2000). The DD may have been formed during the Oligocene and Miocene periods of cooling and dryness that led to the expansion of grasslands and the origin of DD endemic lineages (Simon *et al*, 2009)

Species broadly distributed in the DD frequently present distinct lineages in each biome (Werneck *et al*, 2012a; Fouquet *et al*, 2014), indicating putative isolation events. Diversification in this region has been associated with orogenic events along the Neogene, such as the uplift of the Brazilian Plateau with subsidence of Chaco and other lowlands around 5-7 Ma (Silva, 1995, Carvalho *et al*, 2013) and marine transgressions in the Chacoan region in the Late Miocene (Ruskin *et al*, 2011). Another hypothesis is that Quaternary climate oscillations promoted isolation within the DD, since independent data suggest that during glacial cycles the rain forests expanded through Chaco and Cerrado, scattering and reducing the range of these drier biomes (Nores, 1992; Sobral-Souza *et al*, 2015). If these biomes were primarily affected by orogenic events, species and lineage divergence would date to Neogene (Silva, 1995); however, if glacial cycles' climate changes influenced the biomes more drastically, the taxa would show a Quaternary diversification (Silva, 1995).

Even though there are hypotheses that predict the effects of Quaternary climate changes on Neotropical dry-vegetation dwellers, few studies have effectively tested them. The Pleistocene refugia model (PRM – Vanzolini & Williams, 1981), which was first proposed to explain diversity patterns in the Amazon Forest, asserts that Neotropical rainforests have presented range retraction during glacials with later recovery in interglacials; therefore, open and dry vegetation biomes would have undergone the opposite process, shrinking during wet and warm periods and reaching their maximum geographical distribution during glacials.

Palaeopalynological (Ledru, 2002) and niche modelling (Sobral-Souza *et al*, 2015) studies have agreed that shifts in biome distribution have allegedly occurred over time, but the effects on the demography and genetic structure of the dry-vegetation dwellers remains uncertain.

In the past few years, phylogeography has become an important tool for exploring processes that drove biodiversity evolution in highly diverse regions. For the Neotropics, this approach has been used to test biogeographical hypotheses and make inferences on biome evolution. Despite the recent increase in phylogeographical studies in South America, surveys on open formations are scarce, and just a few studies have focused on organisms endemic to dry biomes (Turchetto-Zolet *et al*, 2013).

Nephila sexpunctata Giebel, 1867 is a spider restricted to southwestern South America, mainly Cerrado, Chaco and adjacent dry forests, and is supposed to have speciated on the continent after a long-distance dispersal event from the Old World in the Miocene/Pliocene (Kuntner *et al*, 2013). It is accepted that *Nephila* species perform aerial dispersal (Lee *et al*, 2015), which may promote connection among distant populations. Spiders have been successfully used as phylogeographical models (e.g. Su *et al*, 2007; Kuntner & Agnarsson, 2011) but in the Neotropical region only two studies have hitherto used this group (Magalhaes *et al*, 2014; Peres *et al*, 2015).

We investigated the recent evolutionary history of Cerrado and Chaco through a phylogeographical study of *Nephila sexpunctata*. Given the poor knowledge about this region's diversification, phylogeographical studies with taxa endemic to these biomes may contribute to clarifying unsolved questions concerning its evolutionary history (Werneck, 2011).

We used mitochondrial and nuclear genetic markers in the first phylogeographical study with an arachnid in this region. We conducted population genetic analyses and estimated divergence times among lineages to detect demographic events. We also used palaeodistribution modelling coupled with a model-based approach to test alternative scenarios for Cerrado and Chaco biogeographic history. Using an approximate Bayesian computation (ABC) framework, we first sought to elucidate when major diversification events took place for *N. sexpunctata*, so we contrasted four feasible scenarios [panmixia (maintenance of gene flow), post-LGM divergence, diversification in Quaternary or an older split in Neogene] to infer what the main factors were that led to the species diversification. Then we aimed to specifically test the predictions made by the PRM, so we contrasted mutually exclusive hypotheses for each scenario: i) *N. sexpunctata* presented no demographic variation in its recent evolutionary history; ii) a previously constant-size population underwent exponential growth; iii) following

the PRM, *N. sexpunctata* presented a demographic expansion from the Last Interglacial (LIG ~120ka) to the Last Glacial Maximum (LGM ~21ka), a colder and drier period, and then experienced a population bottleneck from LGM to present day; iv) LGM conditions were harsh even for dry-vegetation species, so the species suffered a population decrease during the glacial, showing a recovery with exponential growth when the climate became warmer and wetter after LGM. This framework, associated with our ecological niche models and genetic analyses, can help to elucidate historical aspects of DD evolution. We compare the patterns found for *N. sexpunctata* with previous studies using other *Nephila* species as well as studies with co-occurring species. We then proceed to make inferences on the evolution of Cerrado and Chaco based on our results, pinpointing some events that might have guided the organization of diversity as seen in the present.

MATERIALS AND METHODS

Sample collection; DNA extraction, amplification and sequencing

We collected adults of *N. sexpunctata* in most of the species' range, comprising 23 individuals from four locations in the Brazilian Cerrado and 42 from six locations in the Argentinean Chaco and Espinal, herein referred as "Chaco" (Fig 1, Table 1). Opisthosomas and palps were stored in absolute ethanol for species confirmation and deposited in the Coleção Científica de Aracnídeos e Miriápodes of the Instituto Butantan (São Paulo, Brazil), Colección Aracnológica del IADIZA (Mendoza, Argentina) and Museo Argentino de Ciencias Naturales (Buenos Aires, Argentina). Genomic DNA was extracted from 1-8 legs using the Wizard Genomic DNA Purification kit (Promega) following the manufacturer's protocol.

We amplified two mitochondrial regions: for the cytochrome *c* Oxidase subunit I (COI) we used the universal primers for invertebrates, LCO1490 and HCO2198 (Folmer *et al*, 1994).The PCR conditions were: a denaturation step of 3' at 94ºC, 30 cycles of 94ºC for 45", 51ºC for 45" and 72ºC for 2', and a final extension step at 72ºC for 3'. The mitochondrial region comprising the large ribosomal subunit gene 16S rRNA, the tRNA-Leu and the NADH-ubiquinone oxidoreductase (16S-L1-ND1) was amplified using the primers 16S-ar (Palumbi *et al*, 1991) and SPID-ND1 (Hedin, 1997) following the conditions: 94ºC for 5', 28 cycles of 94ºC for 1'30", 53ºC for 1'50" and 72ºC for 2'30", and an extension step at 72ºC for 10'.

We also explored intronic regions of different spider genomes to find candidate markers for phylogeographical studies and developed primers for a novel marker named CHP2 which amplifies a ~300 bp fragment and can be used in other studies on spider phylogeography. Primer

sequences were: CHP2F (GTATCGCAAATGCCTTCAGC) and CHP2R (AACAAAGCAGGCTTCATTCG) and PCR conditions were: 95ºC for 1', 35 cycles of 95ºC for 30", 56ºC for 30", 72ºC for 1' and the final step at 72ºC for 7'. The PCR reactions consisted of 0.5-1µL of template DNA, 0.2µL of 5U Taq DNA Polymerase (Thermo-Fisher), 2-2.8µL of 25mM $MgCl_2$ (Thermo-Fisher), 2.5µL of 10X KCl Buffer (Thermo-Fisher), 0.5-1µL of 10mM dNTP mix (GE Healthcare), 0.5µL of each primer 10µM, and ddH_2O to the volume of 25µL. The amplified products were analyzed in an automatic capillary sequencer ABI PRISM 3700 DNA Analyzer (Applied Biosystems).

Alignment, genetic diversity and haplotype reconstruction

The sequences were aligned in MEGA 7.0 (Kumar *et al*, 2016) and were visually inspected; we found the best fitting substitution model according to AIC on jMODELTEST 2.0 (Darriba *et al*, 2012) and as both mitochondrial regions had the same best substitution model, they were concatenated for all the analyses. For the nuclear markers we used DnaSP (Librado & Rozas, 2009) to infer gametic phases, using a minimum posterior probability threshold of 0.9.

Diversity indices were calculated on ARLEQUIN 3.5 (Excoffier & Lisher, 2010); pairwise genetic distances were estimated on MEGA 7. Relationships among haplotypes were visualized on median-joining networks (Bandelt *et al*, 1999) built on POPART (http://popart.otago.ac.nz).

Population structure and demographic analyses

We assessed population structure with F_{ST} values calculated on ARLEQUIN 3.5. We also performed an analysis of molecular variance (AMOVA) using Cerrado and Chaco as separate groups, to test structure between these biomes. Population structure was also assessed on BAPS 6.0 (Corander *et al*, 2008), which determines the most likely number of clusters (*k*) within a given group of sequences. We allowed *k* to vary between 1 and 20 and performed separate runs for each dataset (mitochondrial and nuclear).

To infer demographic patterns, we performed neutrality tests - Tajima's D (Tajima, 1989), Fu's F_S (Fu, 1997) and R_2 (Ramos-Onsins & Rozas, 2002) - in DnaSP, mismatch distributions analyses in ARLEQUIN, and extended Bayesian skyline plot (EBSP – Heled & Drummond, 2008) in BEAST 1.8.0 (Drummond *et al*, 2012) for the whole dataset and for each mitochondrial group (see Results). Even though these tests have been widely used for

demographic inferences in phylogeographical studies, all of them presume neutrality. In this case, finding similar patterns for mitochondrial and nuclear markers is a stronger evidence for shared demographic changes rather than adaptive selection.

Phylogenetic inferences and divergence times

We conducted a multilocus *BEAST analysis (Heled &Drummond, 2010) in BEAST 1.8.0 in order to estimate divergence among major mitochondrial lineages (according to BAPS) taking into account incomplete lineage sorting. For the mtDNA we used a strict clock (see Results) with a substitution rate of 0.0112 (sd = 0.001) substitutions/site/million years. Bidegaray-Batista & Arnedo (2011) estimated this rate for the Dysderidae family based on the well resolved geochronology of the Mediterranean basin. Recently, Kuntner *et al* (2013) found that substitution rates estimated for the orbicularian families (which include Araneidae) overlapped with those from the Dysderidae, allowing these rates to be implemented to estimate divergence times for orbicularian taxa. For the nuclear marker, we used the phased sequences under a lognormal relaxed clock and estimated the substitution rate in the analysis. As the outgroup we chose *Nephila edulis*, an Australasian species, based on the phylogeny of Kuntner *et al* (2013). COI and 16S-L1-ND1 sequences for this species were obtained from GenBank (accession numbers KC849126 and KC849083, respectively). As no *N. edulis* sequences are available for the newly developed nuclear CHP2 marker, we treated nuclear outgroups sequences as missing data.

A single run was conducted with 200 million generations, sampling every 20000 generations. Convergence with a stationary distribution was checked on TRACER 1.5 (Rambaut *et al*, 2014) through values of Effective Sample Sizes >200 for each prior. We set the burn-in as 10%, discarding the first 1000 trees with TREEANNOTATOR 1.8.0. The resulting species and gene trees were visualized in FIGTREE 1.3.1 (Rambaut, 2012).

Ecological Niche Modelling (ENM)

In the framework applied to model the potential distribution of *N. sexpunctata* under present and past climate scenarios we used the 25 known points of present-day occurrence (Table S1) encompassing our sampling points, data available in SpeciesLink (http://splink.cria.org.br/) and GBIF (www.gbif.org), and points obtained from Instituto Butantan.

Besides the present time, models were projected to mid-Holocene (~6 ka), Last Glacial Maximum (~21ka) and Last Interglacial (~120ka) scenarios. As CCSM3 is the only source for climate data from LIG, we used only this climate database for all temporal scenarios.

To infer potential species distribution, we used five algorithms that represent different modelling techniques (envelope, distance and background) and evaluated the models using the TSS value. We then created a consensus map for each climate scenario (full description of methodology in Appendix S1).

Model Testing

ABC was used for model testing. We performed a hierarchical approach in which we first tested models within four alternative scenarios that represent different phylogeographical hypotheses (Fig. 2) and then compared the best model for each scenario in order to obtain the model with highest posterior probability. This approach provides a reliable framework for testing multiple models (as performed by Peres *et al*, 2015). Scenario 1 (Fig. 2, first line) represents a single panmictic population; scenario 2 (Fig. 2, second line) simulates two populations, one in each biome, with a divergence event dated after the LGM; scenario 3 (Fig. 2, third line) represents the two populations with a split dated in the Quaternary; Scenario 4 (Fig. 2, fourth line) represents the two populations with a Neogene divergence. Within each scenario, four models were tested: i) constant population(s); ii) a constant size-population that experienced demographic expansion at some point in the past; iii) demographic expansion at LIG with a population bottleneck from LGM to the present; iv) a population bottleneck from LIG to LGM, and then a recent post-glacial expansion. For scenarios 2, 3 and 4, migration was also tested. This approach was performed jointly for both datasets. More details on the methodology can be assessed in Appendix S1.

RESULTS

Genetic diversity

The concatenated mitochondrial dataset spanned 1501bp; HKY was the best fitting model for nucleotide substitution. The dataset presented nine haplotypes structured into two groups (Fig. 1) according to BAPS (see *Genetic structure* session). Cerrado populations presented five haplotypes, with four exclusive; Chaco populations also presented five haplotypes, four exclusive.

Diversity levels were very low (Table 2). Three out of ten populations were monomorphic, one in Cerrado and two in Chaco. The CHP2 fragment comprised 232bp and we found seven haplotypes. No indels were found in any of the sequences. TN93+G was the best fitting model for nucleotide substitution. Chaco populations presented all haplotypes while five were detected in Cerrado; only two haplotypes were exclusive, both in Chaco. No population was monomorphic. Diversity levels were also low (Table 3). Nuclear nucleotide diversity was about ten times higher than mitochondrial.

Genetic structure

We found no signs of isolation by distance in either dataset ($p = 0.319$ for nuclear marker; $p = 0.038$ for mitochondrial dataset, but correlation premises were violated – see Fig. S1). The two BAPS mitochondrial groups are separated by only one mutational step. Group I occurs in Cerrado and Group II in Chaco, with some admixture (Fig. 1). The nuclear dataset showed no signs of subdivision in genetic/geographic groups. CHP2 network was highly unstructured (Fig. 1d); five nuclear haplotypes were shared between individuals from both mitochondrial groups.

F_{ST} value was higher for mtDNA (0.59) than for nDNA (0.09), showing a stronger structure. Pairwise values were high in the mitochondrial dataset mainly between populations from different biomes or populations from the same biome belonging to different genetic groups; most of the nuclear comparisons were not significant (Table S4).

AMOVA analyses using biomes as geographical groups (Cerrado x Chaco) were different for each marker; for mtDNA nearly a third of the variation was between groups, evidencing the strong structure between biomes; for the nuclear marker, over 90% of variation is within populations, with no structure pattern (Table 4).

Demographic analyses

Neutrality tests presented non-significant values for mtDNA and nDNA datasets (Tables 2 & 3). The only significant value was Fu's F_S for the population of Parque Nacional del Chaco in the nuclear dataset. Mismatch distribution analyses exhibited contrasting patterns between markers. For the mtDNA, we detected signs of expansion for the whole dataset and for Group II, as evidenced by the non-significant raggedness indexes and unimodal curves of pairwise differences (Fig. 3). The pattern for the nDNA was different, with Group I being the only one presenting evidence of recent expansion (Fig. 3). Multilocus EBSPs pointed to strong and very

recent demographic expansion in all datasets. Given that the generation time for the species is one year, all expansion times were inferred to have taken place in the Holocene, around 10 ka (Fig. 4).

Phylogenetic inferences and divergence times

The maximum likelihood test for molecular clocks implemented in MEGA7 did not reject the strict clock hypothesis for the mitochondrial dataset ($p>0.999$), nor did the path sampling and stepping-stone sampling (Baele *et al*, 2012; Baele *et al*, 2013) implemented in BEAST 1.8.0 (lnBF = 1.92 and 2.58 respectively). The gene trees obtained from the *BEAST analysis (Fig. 5) presented different patterns as the mitochondrial gene tree recovered the same groups inferred by BAPS and the nuclear gene tree presented a less structured, admixed topology. Nuclear diversification was inferred to be older than mitochondrial (0.157 and 0.045 Ma, respectively) but confidence intervals overlapped. The joint species tree analysis estimated the divergence between *N. sexpunctata* and the outgroup (*N. edulis*) around 4.13 Ma [95% of the highest posterior density (HPD) = 3.43-4.90 Ma]. We found high statistical support for the split of the two clusters inferred by BAPS and the divergence between these groups was very recent, estimated around 0.0117Ma (95% HPD = 0.003-0.028).

Ecological Niche Modelling

As TSS values for all algorithms were equal or higher than 0.5 (Table S3), we assumed that they generated reliable predictions. For the LIG scenario there was a very small suitable area for the species' occurrence in the region today covered by Chaco. In the LGM, there were suitable areas in Cerrado, with a distribution disjointed from those in Chaco. Around 6 ka there was a strong expansion of suitable areas in both biomes with the formation of a single joined area of distribution. The current scenario shows an even larger distribution, suggesting recent expansion (Fig. 6).

Model testing

For scenario 1, the most probable model was model 4, which includes a post-LGM expansion after a glacial bottleneck. For scenarios 2 and 3, the best models where those that considered constant populations that experienced exponential growth (models 6 and 10). For scenario 4, the best fit was model 15 (a glacial expansion with a post-LGM bottleneck). For scenarios 3 and 4, models that considered migration were strongly supported (Table 5).

When the analysis was performed among scenarios to find the best overall model, the panmixia (model 4) was preferred, but the post-LGM divergence (model 6) also had a high posterior probability (Table 5). Quaternary and Neogene divergence were least supported. Parameter estimation for models 4 and 6 can be assessed in Appendix S2 (Table S5). Therefore, the results endorse a recent demographic expansion.

DISCUSSION

Phylogeographical patterns of N. sexpunctata

The ABC results showed as the best overall model one single population that experienced a bottleneck during the glacial with a post-LGM recovery, opposite to what was expected by the PRM. The fact that the panmixia model was the preferred is against our assumption that different lineages were isolated in each biome, as found for lizards and frogs (Werneck *et al*, 2012a; Recoder *et al*, 2014), but we could not discard a post-LGM divergence as model 6 also showed a high posterior probability. The significant genetic structure among *N. sexpunctata* populations coupled with the low genetic divergence among haplotypes and the strong dispersal in the species may have generated genetic patterns that fit similarly the two scenarios. In both cases, the results endorsed a scenario of recent demographic expansion, and against what was expected by the PRM.

The low genetic diversity of *N. sexpunctata* contrasts with the values reported for other *Nephila* species (Su *et al*, 2007; Su *et al*, 2011). Low levels of genetic diversity are common for Cerrado and Chaco species (Babb *et al*, 2011; Brito *et al*, 2013) and are frequently associated with bottlenecks due to the reduction of the suitable habitats of species adapted to dry conditions during the Quaternary (e.g. Bonatelli *et al*, 2014).

Discrepancies between nuclear and mitochondrial structuring patterns have been recurrently reported (Turchetto-Zolet *et al*, 2013) and are a likely product of multiple factors such as mtDNA introgression, incomplete lineage sorting, demographic asymmetries (Toews & Brelsford, 2012) and potential selection on the mitochondrial genome (Bazin *et al*, 2006). A slower rate of evolution, coupled with an effective population size four times higher for the nuclear genome compared to the mitochondrial genome, may also lead to the distinct structuring patterns observed. In *N. sexpunctata*, as for many other terrestrial invertebrates, these factors might help explaining the higher structure found on the mitochondrial markers.

The mitochondrial dataset presented two main lineages almost entirely related to each biome (Cerrado/Chaco, Fig. 1), and indeed a significant part of the species' diversity is organized between biomes (Table 4). As BAPS assigns groups for both genetic and geographic similarities, geography may play an important role in the characterization of the two clusters, since just one mutational step separated them. The admixture observed - represented by the occurrence of both mitochondrial lineages in some populations - may be due to the high dispersal ability of *Nephila* species (Kuntner & Agnarsson, 2011).

The signs of recent demographic expansion in *N. sexpunctata* after LGM (Figs. 3 and 4) are congruent with the post-glacial habitat expansion inferred by ENMs and the results from ABC. Lack of significance for neutrality tests may be due to the very low genetic variability in our sample, leading to analyses' low statistical power (Excoffier *et al*, 2009). Recent population and range expansions have been found for several DD species (e.g. Prado *et al*, 2012; Novaes *et al*, 2013), and are related to habitat expansion after the retraction of Cerrado vegetation during LGM (Ledru *et al*, 1996) due to extreme arid and cold conditions. Hence, the *N. sexpunctata* range would be reduced in LGM, expanding in the Holocene as climate conditions in southern Cerrado became warmer and wetter (Ledru, 2002), more suitable for the species' occurrence. These results do not support the idea of interglacial refuges for dry-habitat adapted organisms as suggested for cactus species (Bonatelli *et al*, 2014) and indicate that species with different ecological requirements may respond differently to climate fluctuations, with distinct factors shaping each species distribution (Prado *et al*, 2012).

N. sexpunctata divergence from outgroup *N. edulis* was estimated on Pliocene, supporting speciation after a recent dispersal event of an Australasian *Nephila* lineage to the Neotropics, as suggested by Kuntner *et al* (2013). Studies using molecular dating have shown that dispersal events between the Palaeotropics and the Neotropics were frequent in the past few million years and are a good explanation for diversity patterns observed in many plants and animals (e.g. Blaimer, 2012; Christenhusz & Chase, 2013). Therefore, an open vegetation-adapted *Nephila* lineage, ancestral of *N. sexpunctata,* would have arrived in the Neotropics after the settlement of the DD, where it established itself. Diversification within the species, on the other hand, is very recent; the coalescence of all *N. sexpunctata* sequences was estimated around 11.7 ka, indicating that the extant diversity radiated during the Quaternary. This pattern is consistent with simulations for Cerrado range during the Quaternary, which suggests unstable distribution throughout the last glacial cycle (Terribile *et al*, 2012; Werneck *et al*, 2012b); these

data combined support the hypothesis that climate shifts may have played an important role in the diversification of Neotropical dry biomes (Silva, 1995).

Effects of Quaternary climate fluctuations on DD diversification

Despite the growing interest in Neotropical open formations, there is still a scarcity of testable hypotheses concerning DD evolution. Whereas a substantial amount of work concerning rain forests has indicated that geomorphologic (Hoorn *et al*, 2010, Thomé *et al*, 2014) and climate (Cheng *et al*, 2013; Sobral-Souza *et al*, 2015) events may have worked together to shape biodiversity evolution in the Amazon and Atlantic Forest, the evolution of DD organisms is frequently explained by geologic events such as the uplift of the Brazilian plateau (Werneck, 2011; Carvalho *et al*, 2013). Indeed, recent works have highlighted the role of Pleistocene climate oscillations in shaping current genetic variation, mainly through shifts in Cerrado range that caused changes in population size and bottlenecks (Diniz-Filho *et al*, 2016).

Ecological Niche Modelling for different DD biomes (Werneck *et al*, 2011; Terribile *et al* 2012; Werneck *et al*, 2012b) does not support the predictions of the PRM, which include range expansion of open formations towards rain forests during glacial periods. The models have shown that conditions were probably too cold and dry for large extents of these biomes during LGM, and they presumably presented smaller ranges in this period with later expansion in the Holocene/present-day. Genetic and paleodistributional data for *N. sexpunctata* show a post-LGM demographic and spatial expansion pattern that fits a scenario of post-glacial expansion of South American dry biomes. During LGM, the retraction of Chaco and Cerrado could have decreased habitat suitability for *N. sexpunctata* occurrence, leading to population bottlenecks (as reinforced by ABC results), causing the low levels of genetic diversity observed in the species nowadays. In a phylogeographical study with *Sicarius cariri*, a spider endemic to another counterpart of the DD (Caatinga), Magalhaes *et al* (2014) found evidence of demographic retraction during glacials with a Holocene recovery, in a pattern similar to what was found for *N. sexpunctata* and against that predicted by the PRM. This finding suggests that glaciations might have affected some DD species in a different way from what was expected by the PRM, possibly because of common mechanisms related to habitat retraction due to harsh climate conditions.

Even though the effects of savanna expansion towards moist forests have long been acknowledged (Brown & Ab'Saber, 1979), the role of rain forest invasions in the diversification of DD organisms has only recently been recognized. Expansion of the Amazon and Atlantic

Forest's spatial ranges seems to have shaped the genetic diversity in *Pleurodema*, a genus of Caatinga-endemic frogs, by creating intermittent forest corridors that promoted recurrent isolation among populations (Thomé *et al*, 2016). Likewise, expansion of the rain forests towards Cerrado and Chaco (Nores, 1992; Sobral-Souza *et al*, 2015) may have contributed to the structuring observed mainly in the mitochondrial dataset, the low genetic diversity indices and the demographic changes inferred.

In conclusion, our work reinforces the complexity of the diversification process in the Neotropical region. We demonstrated that Quaternary climate oscillations have affected Cerrado and Chaco, and consequently the diversification of DD organisms. This has to be taken into account, besides the frequently invoked main geological events. Our results also show that, unlike what was predicted by the PRM, some open-vegetation organisms have presented recent expansion, indicating that this period was too cold and dry even for their occurrence. The extent to which organisms from other Neotropical dry biomes were affected in a similar manner remains to be elucidated.

ACKNOWLEDGEMENTS

We thank ICMBio for collection permits (#27127 and 38889), Douglas de Araujo for donation of samples, the Coleção de Invertebrados em Meio Líquido da Universidade Federal de Minas Gerais, especially Dr. Adalberto José dos Santos, for museum samples, and CONICET for help during Argentina fieldwork. Manolo Perez and Isabel Bonatelli helped with ABC. Jair F. Mendes Jr, Maria Florencia F. Campón, Susana Lagos and Gualberto Zalazar helped with fieldwork.

We also thank the following entities for financial support: Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (fellowship #141040/2013-0 to LFMB); São Paulo Research Foundation – FAPESP (grants #2012/02526-7, #2013/50491-0 and #2013/08293-7).

REFERENCES

Babb, P.L., Fernandez-Duque, E., Baiduc, C.A., Gagneux, P., Evans, S. & Schurr, T.G. (2011) mtDNA diversity in Azara's Owl Monkeys (*Aotus azarai azarai*) of the Argentinean Chaco. *American Journal of Physical Anthropology*, **146**, 209-224.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A. & Alekseyenko, A.V. (2012) Improving the accuracy of demographic and molecular clock model comparison while

accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, **29**, 2157-2167.

Baele, G., Li, W.L.S., Drummond, A.J., Suchard, M.A. & Lemey, P. (2013) Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, **30**, 239-243.

Bandelt, H.J., Forster, P. & Rohl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Bazin, E., Glémin, S. & Galtier, N. (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science*, **312,** 570–572.

Bidegaray-Batista, L. & Arnedo, M.A. (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology*, **11**, 317.

Blaimer, B.B. (2012) Acrobat ants go global – Origin, evolution and systematic of the genus Crematogaster (Hymenoptera:Formicidae). *Molecular Phylogenetics and Evolution*, **65**(2), 421-436.

Bonatelli, I.A.S., Perez, M.F., Peterson, A.T., Taylor, N.P., Zappi, D.C., Machado, M.C., Koch, I., Pires, A.H.C. & Moraes, E.M. (2014) Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, **23**, 3044–3063.

Brito, R.M., Francisco, F.O., Françoso, E., Santiago, L.R. & Arias, M.C. (2013) Very low mitochondrial variability in a stingless bee endemic to cerrado. *Genetics and Molecular Biology*, **36**(1), 124-128.

Brown, K.S. & Ab'Saber, A.N. (1979) Ice-age forest refuges and evolution in Neotropics: correlation of paleoclimatoligical, geomorphological and pedological data with biological endemism. *Paleoclimas*, **5**, 1–30.

Carvalho, A.L., de Britto, M.R. & Fernandes, D.S. (2013) Biogeography of the lizard genus *Tropidurus* Wied-Neuwied, 1825 (Squamata:Tropiduridae): distribution, endemism, and area relationships in South America. *PLoS ONE*, **8**, e59736.

Cheng, H., Sinha, A., Cruz, F.W., Wang, X., Edwards, R.L., d'Horta, F.M., Ribas, C.C., Vuille, M., Scott, L.D. & Auler, A.S. (2013) Climate change patterns in Amazonia and biodiversity. *Nature Communications*, **4**, 1411.

Christenhusz, M.J.M. & Chase, M.W. (2013) Biogeographical patterns of plants in the Neotropics: dispersal rather than plate tectonics is most explanatory. *Botanical Journal of the Linnean Society*, **171**, 277–286.

Corander, J., Sirén, J. & Arjas, E. (2008) Bayesian Spatial Modelling of Genetic Population Structure. *Computational Statistics*, **23**, 111-129.

Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9(8)**, 772.

Diniz-Filho, J.A.F., Barbosa, A.C.O.F., Collevatti, R G., Chaves, L.J., Terribile, L.C., Lima-Ribeiro, M.S. & Telles, M.P.C. (2016) Spatial autocorrelation analysis and ecological niche modelling allows inference of range dynamics driving the population genetic structure of a Neotropical savanna tree. *Journal of Biogeography*, **43**: 167–177.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology And Evolution*, **29**: 1969-1973

Excoffier, L., Foll, M. & Petit, R.J. (2009) Genetic consequences of range expansions. *Annual Review in Ecology, Evolution and Systematics*, **40**, 481–501.

Excoffier, L. & Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**(5), 294–299.

Fouquet, A., Cassini, C.S., Haddad, C.F.B., Pech, N. & Rodrigues, M.F. (2014) Species delimitation, patterns of diversifications and historical biogeography of the Neotropical frog genus Adenomera (Anura:Leptodactylidae). *Journal of Biogeography*, **41**, 855-870.

Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Hedin, M.C. (1997) Molecular phylogenetics at the population/species interface in cave spiders of the southern Appalachians (Araneae:Nesticidae:Nesticus). *Molecular Biology and Evolution*, **14**, 309–324.

Heled, J. & Drummond, A.J. (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.

Heled, J. & Drummond, A.J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T. & Antonelli, A. (2010) Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927-931.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870-1874.

Kuntner, M. & Agnarsson, I. (2011) Phylogeography of a successful aerial disperser: the golden orb spider *Nephila* on Indian Ocean islands. *BMC Evolutionary Biology.* **11**:119.

Kuntner, M., Arnedo, M.A., Trontelj, P., Lokovsek, T. & Agnarsson, I. (2013) A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution*, **69**, 961-979.

Ledru, M.P., Braga, P.I.S., Soubiès, F., Fournier, M., Martin, L., Kenitiro, S. & Turcq, B. (1996) The last 50,000 years in the Neotropics (Southern Brazil): evolution of vegetation and climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **123**: 239–257.

Ledru, M.P. (2002) Late quaternary history and evolution of the Cerrados as revealed by palynological records. *The cerrados of Brazil: ecology and natural history of a neotropical savanna*. (ed by P.S. Oliveira and R.J. Marquis), pp. 33–68. Columbia University Press, New York.

Lee, V.M.J., Kuntner, M. & Li, D. (2015) Ballooning behavior in the golden orbweb *spider Nephila pilipes* (Araneae:Nephilidae). *Frontiers in ecology and evolution*, **3**, 2:1-5.

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Magalhaes, I.L., Oliveira, U., Santos, F.R., Vidigal, T.H., Brescovit, A.D. & Santos, A.J. (2014) Strong spatial structure, Pliocene diversification and cryptic diversity in the Neotropical dry forest spider *Sicarius cariri*. *Molecular Ecology*, **23**, 5323–5336.

Morley, R.J. (2000) *Origin and evolution of tropical rain forests*. John Wiley & Sons Ltd, Chinchester.

Nores, M. (1992) Bird speciation in subtropical South America in relation to forest expansion and retraction. *The Auk*, **109**, 346-357.

Novaes, R.M.L., Ribeiro, R.A., Lemos-Filho, J.P. & Lovato, M.B. (2013) Concordance between Phylogeographical and Biogeographical Patterns in the Brazilian Cerrado:

Diversification of the Endemic Tree *Dalbergia miscolobium* (Fabaceae). *PLoS ONE* **8**(12): e82198.

Palumbi, S., Martin, A., Romano, S., McMillan, W.O., Stice, L. & Grabowski, G. (1991) The Simple Fool's Guide to PCR. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, USA.

Peres, E.A., Sobral-Souza, T., Perez, M.F., Bonatelli, I.A.S., Silva, D.P., Silva, M.J. & Solferini, V.N. (2015) Pleistocene Niche Stability and Lineage Diversification in the Subtropical Spider *Araneus omnicolor* (Araneidae). *PLoS ONE*, **10**(4): e0121543.

Prado, C.P.A., Haddad, C.F.B. & Zamudio, K.R. (2012) Cryptic lineages and Pleistocene population expansion in a Brazilian Cerrado frog. *Molecular Ecology*, **21**, 921-941.

Rambaut, A. (2012) *FigTree, version 1.4*, available at http://tree.bio.ed.ac.uk/software/figtree.

Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A.J. (2014) *Tracer, version 1.6*, Available at http://beast.bio.ed.ac.uk/Tracer

Ramos-Onsins, S.E. & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Recoder, R.S., Werneck, F., Teixeira, M., Colli, G.R., Sites, J.W. & Rodrigues, M.T. (2014) Geographic variation and systematic review of the lizard genus *Vanzosaura* (Squamata, Gymnophthalmidae), with the description of a new species. *Zoological Journal of the Linnean Society*, **171**: 206–225.

Ruskin, B. G., Dávila, F. M., Hoke, G. D., Jordan, T. E., Astini, R. A. & Alonso, R. (2011) Stable isotope composition of middle Miocene carbonates of the Frontal Cordillera and Sierras Pampeanas: did the Paranaense seaway flood western and central Argentina? *Palaeogeography, Palaeoclimatology. Palaeocology.* **308**, 293–303.

Silva, J.M.C. (1995) Biogeographic analysis of the South American Cerrado avifauna. *Steenstrupia*, **21**, 49-67.

Simon, M.F., Grether, R., de Queiroz, L.P., Skema, C., Pennington, R.T. & Hughes, C.E. (2009) Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences USA*, **106**, 20359–20364

Sobral-Souza, T., Lima-Ribeiro, M.S. & Solferini, V.N. (2015) Biogeography of Neotropical Rainforests: past connections between Amazon and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, **29**, 643–655.

Su, Y.-C., Chang, Y.-H., Lee, S.-C. & Tso, I.-M. (2007) Phylogeography of the giant wood spider (*Nephila pilipes*, Araneae) from Asian–Australian regions. *Journal of Biogeography*, **34,** 177–191.

Su, Y-C., Chang, Y-H., Smith, D., Zhu, M-S., Kuntner, M. & Tso, I-M. (2011) Biogeography and speciation patterns of the golden orb spider genus *Nephila* (Araneae:Nephilidae) in Asia. *Zoological Science*, **28**(1), 47-55.

Tajima, F. (1989) Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Terribile, L.C., Lima-Ribeiro, M.S., Araújo, M., Bizao, N., Collevatti, R.G., Dobrovolski, R., Franco, A., Guilhaumon, F., Lima, J.S., Murakami, D.M., Nabout, J.C., Oliveira, G., Oliveira, L.K., Rabello, S.G., Rangel, T.F., Simon, L.M., Soares, T.N., Telles, M.P.C. & Diniz-Filho, J.A.F. (2012) Areas of climate stability of species ranges in the Brazilian Cerrado: disentangling uncertainties through time. *Natureza & Conservacao*, **10**, 152–159.

Thomé, M.T.C., Zamudio, K.R., Haddad, C.F. & Alexandrino, J. (2014) Barriers, rather than refugia, underlie the origin of diversity in toads endemic to the Brazilian Atlantic Forest. *Molecular Ecology*, **23**, 6152–6164.

Thomé, M.T.C., Sequeira, F., Brusquetti, F., Carstens, B., Haddad, C.F.B., Rodrigues, M.T. & Alexandrino, J. (2016) Recurrent connections between Amazon and Atlantic forests shaped diversity in Caatinga four-eyed frogs. *Journal of Biogeography*, **43**(5), 1045-1056.

Toews, D.P.L. & Brelsford, A. (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Turchetto-Zolet, A.C., Pinheiro, F., Salgueiro, F. & Palma-Silva, C. (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Vanzolini, P.E. & Williams, E.E. (1981) The vanishing refuge: A mechanism for ecogeographic spectiation. *Papéis avulsos de Zoologia*, **34(23)**, 251-255.

Werneck, F.P. (2011) The diversification of eastern South American open vegetation biomes: historical biogeography and perspectives. *Quaternary Science Reviews*, **30**, 1630–1648.

Werneck, F.P., Costa, G.C., Colli, G.R., Prado, D. & Sites Jr., J.W. (2011) Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecology and Biogeography*, **20**, 272-288.

Werneck, F.P., Gamble, T., Colli, G.R., Rodrigues, M.T. & Sites Jr., J.W. (2012) Deep diversification and long term persistence in the south American 'dry diagonal': integrating continent-wide phylogeography and distribution modeling of geckos. *Evolution*, **66**, 3014–3034.

Werneck, F.P., Nogueira, C., Colli, G.R., Sites Jr, J.W. & Costa, G.C. (2012) Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *Journal of Biogeography*, **39**, 1695–1706.

SUPPLEMENTARY INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Details on methodology.

Appendix S2 Supplementary tables.

Appendix S3 Supplementary figures.

DATA ACCESSIBILITY

All sequences are available on GenBank (accession numbers – COI: KX028886 - KX028950; 16S-L1-NAD1: KX028951 - KX029015; CHP2: KX029016 - KX029080). Xml input files used in Bayesian analyses and python scripts for coalescent simulations are available on GitHub (https://github.com/luizbartoleti/Nsexpunctata).

BIOSKETCH

Luiz Filipe de Macedo Bartoleti is interested in the phylogeography of Neotropical organisms, with emphasis on the evolutionary history of Neotropical biomes revealed by widely distributed species. All authors are interested in evolutionary biology and biogeography.

Author Contributions:

LFMB and VNS conceived and designed the study. LFMB, EAP and FVHMF carried out fieldwork. LFMB and MJS performed the experiments. MJS and VNS contributed with reagents and material tools. LFMB, EAP, TSS, FVHMF and VNS analyzed the data. LFMB, EAP, TSS and VNS wrote the paper. All authors have read and approved the final version submitted.

Editor: Alain Vanderpoorten

TABLES

Table 1: List of *Nephila sexpunctata* sampling locations with details on biome and geospatial coordinates.

Location	Biome	Latitude	Longitude
Brazil – Tupã (TP)	Cerrado	21°55'01.8"S	50°30'04.6"W
Brazil – Pirenópolis (PI)	Cerrado	15°55'24.2"S	49°12'50.6"W
Brazil – Catalão (CT)	Cerrado	18°06'50.0"S	47°37'16.0"W
Brazil – Campo Grande (CG)	Cerrado	21°24'13.8"S	53°40'43.8"W
Argentina – Paraná (PA)	Espinal	31°43'32.1"S	60°17'58.1"W
Argentina – Federal (FE)	Espinal	30°56'01.0"S	58°43'44.9"W
Argentina – Mercedes (ME)	Espinal	29°05'44.6"S	58°19'46.5"W
Argentina – Santa Maria (SM)	Chaco	28°02'29.6"S	58°06'02.2"W
Argentina – Pq. Nacional del Chaco (CH)	Chaco	26°48'31.1"S	59°36'21.5"W
Argentina – Basail (BA)	Chaco	27°52'01.7"S	59°17'22.0"W

Table 2: Mitochondrial genetic diversity indices and neutrality tests for *Nephila sexpunctata* populations and genetic groups inferred by BAPS. N = number of sequences; h = number of haplotypes; ss = number of segregating sites; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; D = Tajima's D; R2 = R2 statistic; F_{ST} = fixation index, sd = standard deviation. * $p < 0.05$; ** $p < 0.02$.

Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**	**R2**	**F_{ST}**
TP	8	2	1	0.536 (0.123)	0.00036 (0.00037)	0.866	1.166	0.268	-
PI	8	3	2	0.607 (0.164)	0.00045 (0.00044)	-0.478	-0.448	0.197	-
CT	3	1	0	0.000 (0.000)	0.00000 (0.00000)	-	0.000	-	-
CG	4	2	2	0.500 (0.265)	0.00067 (0.00066)	1.099	-0.710	0.433	-
PA	3	1	0	0.000 (0.000)	0.00000 (0.00000)	-	0.000	-	-
FE	8	2	1	0.250 (0.180)	0.00017 (0.00024)	-0.182	-1.055	0.331	-
ME	8	2	2	0.429 (0.169)	0.00057 (0.00051)	1.653	0.414	0.214	-
SM	8	1	0	0.000 (0.000)	0.00000 (0.00000)	-	0.000	-	-
CH	7	4	4	0.809 (0.130)	0.00089 (0.00071)	-0.914	-0.876	0.192	-
BA	8	2	1	0.536 (0.536)	0.00036 (0.00037)	0.866	1.166	0.268	-
Group I	32	4	3	0.599 (0.049)	0.00045 (0.00039)	-0.451	-0.197	0.118	-
Group II	33	5	4	0.481 (0.089)	0.00035 (0.00033)	-2.252	-1.166	0.084	-
Total	65	9	8	0.773 (0.027)	0.00081 (0.00058)	-2.771	-0.732	0.074	0.59

Table 3: Nuclear genetic diversity indices and neutrality tests for *Nephila sexpunctata* populations and genetic groups inferred by BAPS. N = number of sequences; h = number of haplotypes; ss = number of segregating sites; Hd = haplotype diversity; π = nucleotide diversity; F_S = Fu's F_S; D = Tajima's D; R2 = R2 statistic; F_{ST} = fixation index, sd = standard deviation. * $p < 0.05$; ** $p < 0.02$.

Location	**N**	**h**	**ss**	**Hd (sd)**	**π (sd)**	**F_S**	**D**	**R2**	**F_{ST}**
TP	16	4	3	0.517 (0.132)	0.00295 (0.0026)	-1.098	-0.708	0.123	-
PI	16	5	3	0.608 (0.130)	0.00366 (0.0031)	-1.845	-0.173	0.145	-
CT	6	3	3	0.733 (0.155)	0.00546 (0.0046)	0.209	-0.185	0.268	-
CG	8	3	3	0.678 (0.122)	0.00570 (0.0045)	0.723	-0.585	0.210	-
PA	6	2	2	0.333 (0.215)	0.00287 (0.0029)	0.952	-1.132	0.373	-
FE	16	4	3	0.617 (0.096)	0.00305 (0.0027)	-1.019	-0.628	0.134	-
ME	16	5	4	0.808 (0.064)	0.00578 (0.0042)	-0.689	0.353	0.168	-
SM	16	4	3	0.642 (0.081)	0.00323 (0.0028)	-0.893	-0.494	0.138	-
CH	14	5	3	0.670 (0.126)	0.00384 (0.0032)	-1.933**	-0.173	0.148	-
BA	16	2	1	0.533 (0.046)	0.00230 (0.0023)	1.362	1.529	0.267	-
Group I	64	6	4	0.696 (0.044)	0.00467 (0.0034)	-0.542	0.603	0.135	-
Group II	66	6	4	0.602 (0.040)	0.00311 (0.0026)	-1.754	-0.300	0.092	-
Total	130	7	5	0.653 (0.031)	0.00399 (0.0031)	-1.217	0.014	0.093	0.09

Table 4: Results of the Analysis of Molecular Variance (AMOVA) for mitochondrial and nuclear datasets for *Nephila sexpunctata* considering biomes (Cerrado and Chaco) as groups. d.f. = degrees of freedom.

Source of variation	**d. f.**	**Variation (%)**	**p-value**
mtDNA			
Among biomes	1	29.07	p = 0.047
Among populations within biomes	8	33.14	p < 0.001
Within populations	55	37.79	p < 0.001
nDNA			
Among biomes	1	2.13	p = 0.171
Among populations within biomes	8	6.36	p = 0.011
Within populations	120	91.52	p = 0.006

Table 5: Results of the model selection performed for both datasets. For every scenario posterior probabilities (PP) of each model are displayed, including the most recurrently recovered model (in bold). The analyses were performed within and among scenarios. Details on each model can be found in Appendix S1.

		Within scenarios		**Among scenarios**
Scenario	**Model**	**PP**	**Migration (Y/N) – PP***	**PP**
1 (panmixia)	**1**	0.217	Not tested	-
	2	0.308		-
	3	0.050		-
	4	**0.428**		**0.489**
2 (post-LGM divergence)	**5**	0.307	(N) 0.503	-
	6	**0.499**		0.412
	7	0.046		-
	8	0.148		-
3 (Quaternary divergence)	**9**	0.300	(Y) 0.663	-
	10	**0.571**		0.096
	11	0.107		-
	12	0.022		-
4 (Neogene divergence)	**13**	0.006	(Y) 0.999	-
	14	0.002		-
	15	**0.973**		0.003
	16	0.019		-

*For each scenario, we chose the best model and then contrasted it with a similar one that considered bidirectional migration, and performed a new selection between these two models. In this column, we show if the model with (Y) or without (N) migration was preferred, and the PP of this model. Panmixia models were not tested for migration as they assume a single population with random mating.

FIGURE LEGENDS

Figure 1: a) Map with *Nephila sexpunctata* sampling locations showing four sites in Brazilian Cerrado (dark gray) and six in Argentinean Dry Forests (Chaco and Espinal - lighter grays; Espinal populations are considered in the Chaco dataset). Pie charts for each population represent proportion of mitochondrial haplotypes (c) in given population. b) BAPS chart showing individual assignment to mitochondrial groups represented by different colors – Group I: red, Group II: blue. c) Mitochondrial network (COI + 16S-L1-ND1) highlighting nine haplotypes divided into two mitochondrial groups inferred by BAPS: Group I occurs predominantly in Cerrado and Group II in Chaco. Circle size is proportional to frequency, and dashes represent mutational events. d) Nuclear network (CHP2) coloured by the occurrence of each haplotype in Chaco or Cerrado.

Figure 2: Demographic scenarios hypothesized for *Nephila sexpunctata* evolution in Cerrado and Chaco. Scenario 1: panmictic population; scenario 2: two populations, one in each biome, with a post-LGM divergence; scenario 3: two populations with a Quaternary divergence; scenario 4: two populations with a Neogene split. Within each scenario, four hypotheses were tested: i) no demographic changes along time (models 1, 5, 9 and 13); ii) exponential growth of a previous constant-size population (models 2, 6, 10 and 14); iii) glacial expansion followed by interglacial bottlenecks (as predicted by the PRM, models 3, 7, 11 and 15); iv) glacial bottlenecks with post-glacial recovery (models 4, 8, 12 and 16). The best overall model is highlighted. LGM = Last Glacial Maximum (~21ka); LIG = Last Interglacial (~120ka).

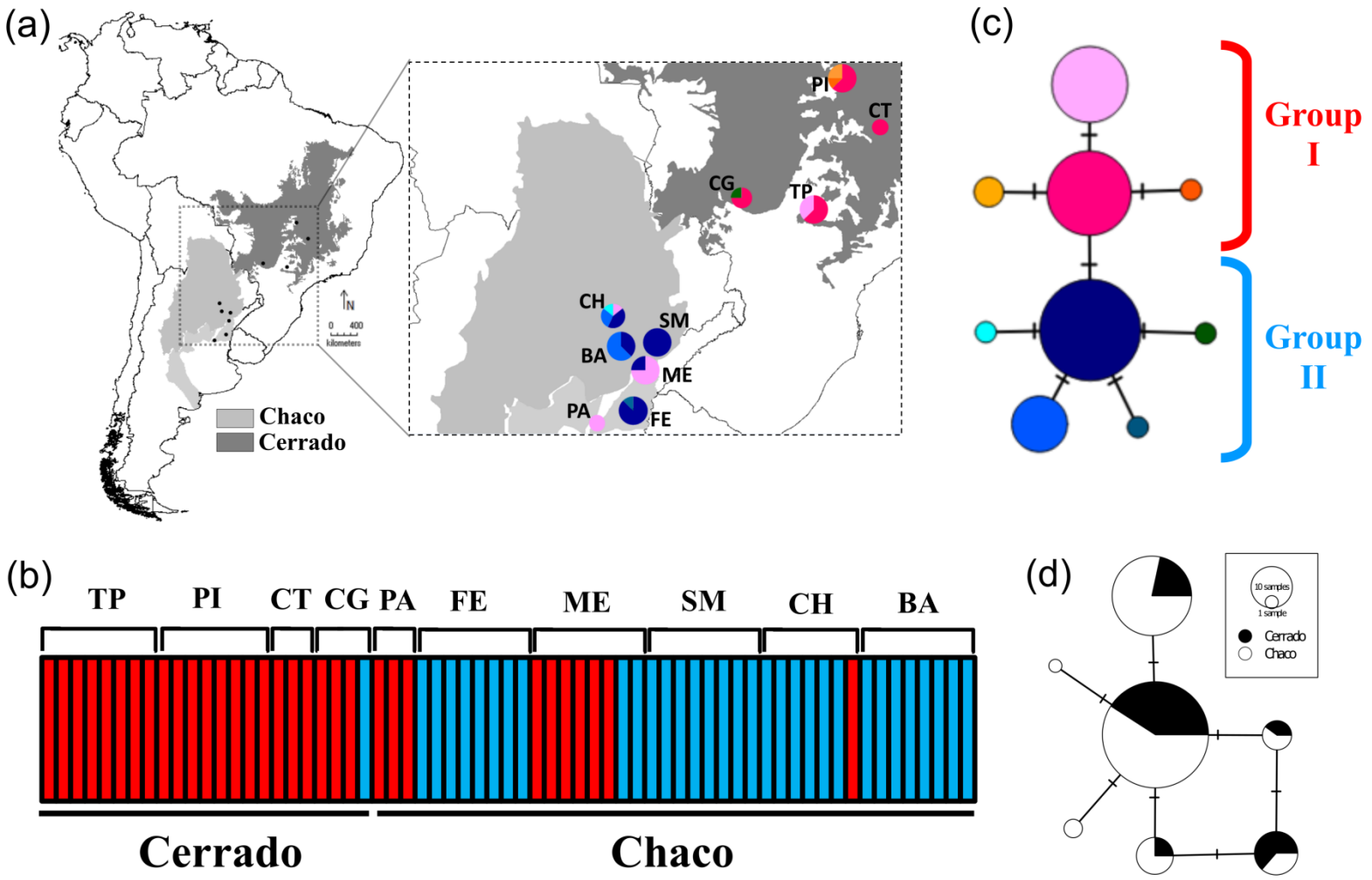
Figure 3: Results of mismatch distribution analyses performed for *Nephila sexpunctata* sequences. mtDNA (above) and nDNA (below) were analyzed separately. Results were obtained for all sequences for each marker, and for each mitochondrial group separately.

Columns represent the observed pairwise comparisons between individuals, and the continuous line represents the expected comparison in a scenario of sudden expansion. r = Harpending's Raggedness index; p = p-value.

Figure 4: Results of Extended Bayesian Skyline Plots performed for *Nephila sexpunctata*. Analyses were performed using both markers jointly. Continuous line represents the median, and the gray area represents 95% HPD. a) Results for the whole dataset, b) Result obtained only for Group I individuals. c) Result obtained only for Group II individuals.

Figure 5: Bayesian gene trees inferred for *Nephila sexpunctata* mitochondrial (above) and nuclear (below) datasets. Branch colors represent the two mitochondrial groups inferred by BAPS. Black squares represent nodes with posterior probability > 0.95.

Figure 6: Ecological Niche Modelling showing suitable areas for *Nephila sexpunctata* occurrence over the last glacial cycle. Top left – 120ka; Top right – 21ka; Bottom left – 6ka; Bottom right – present day.


(a)
PI
CT
CG
TP
CH
SM
BA
ME
PA
FE
Chaco
Cerrado
(b)
TP
PI
CT
CG
PA
FE
ME
SM
CH
BA
Cerrado
Chaco
(c)
Group I
Group II
(d)
10 samples
1 sample
Cerrado
Chaco

1

2 **Figure 1**

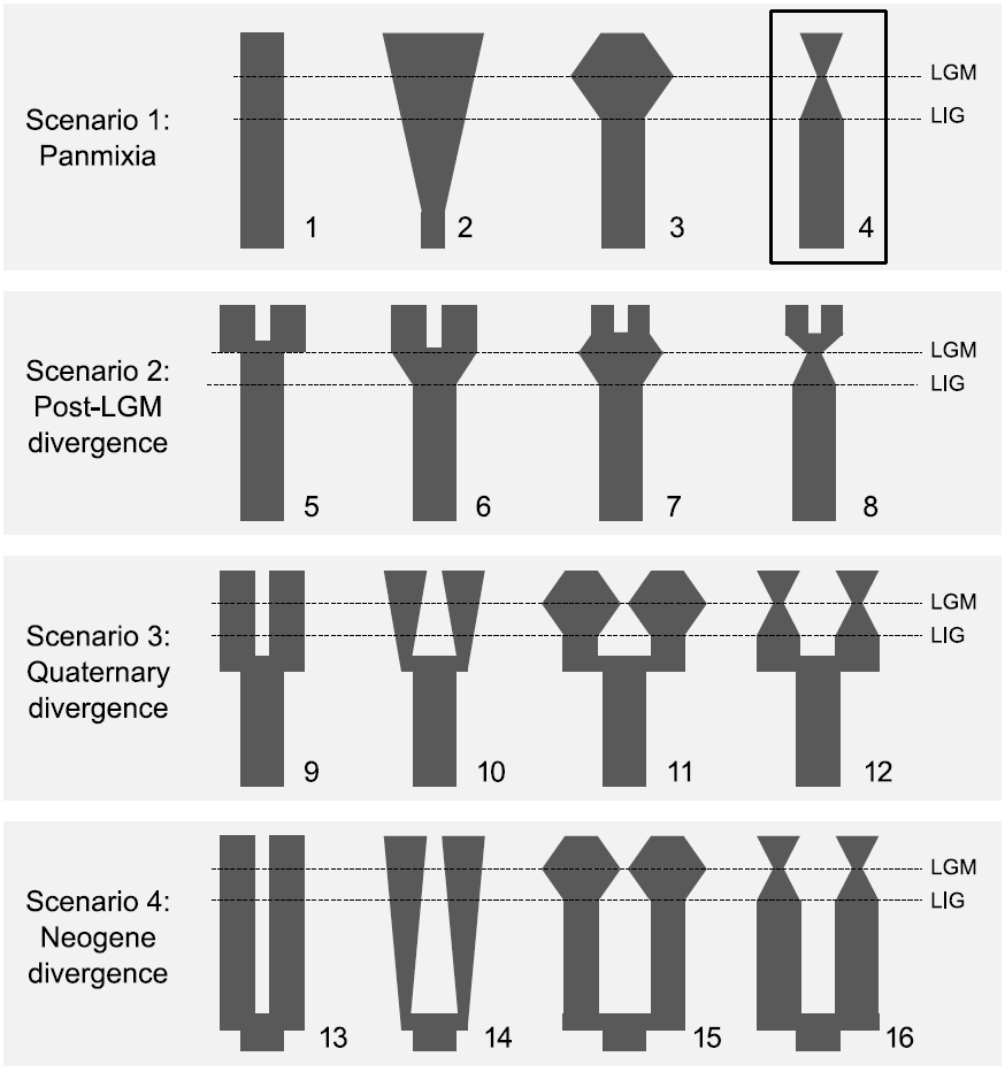
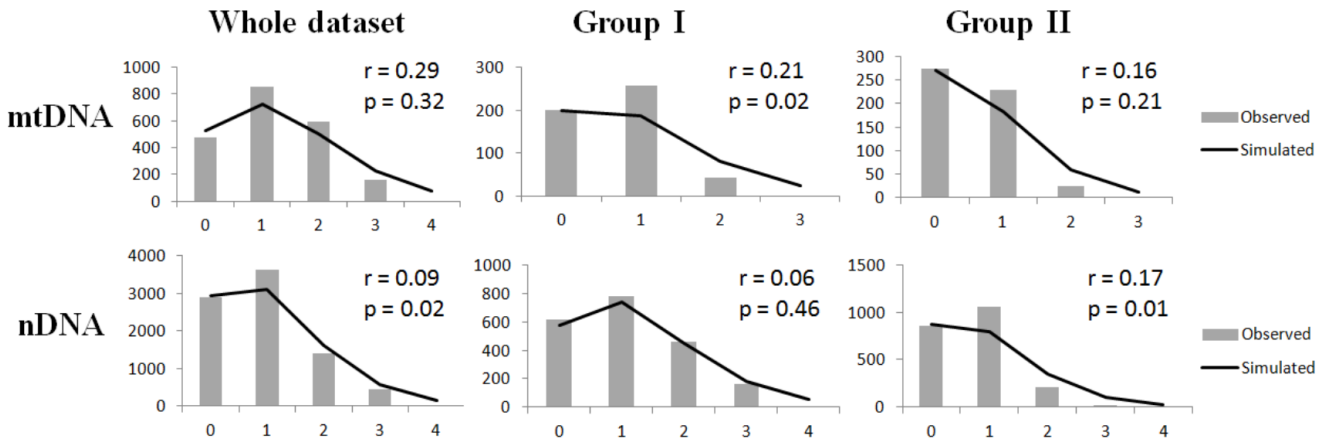
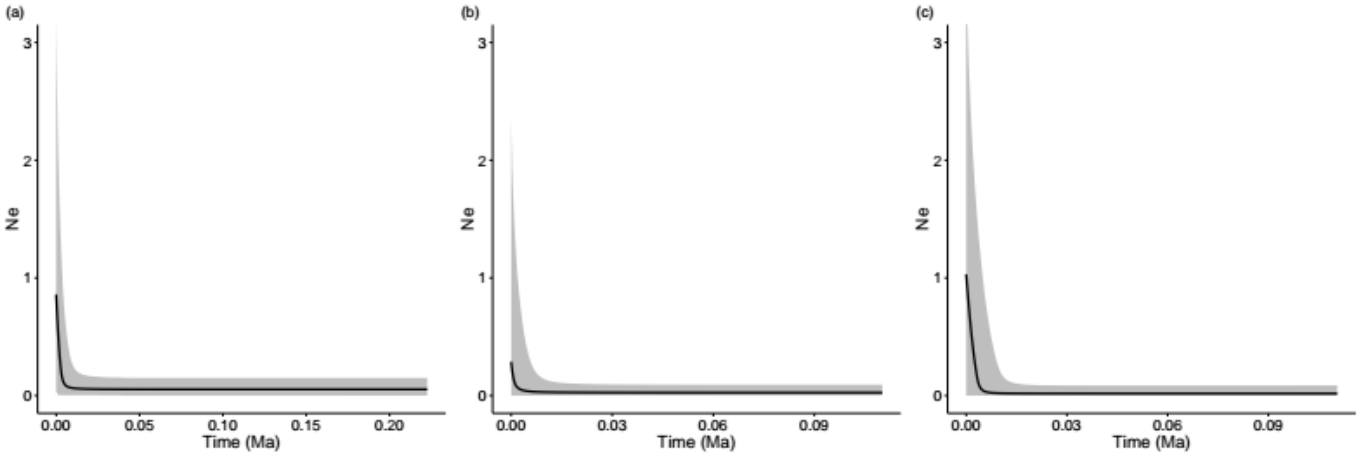

Scenario 1:
Panmixia
LGM
LIG
1
2
3
4
Scenario 2:
Post-LGM
divergence
LGM
LIG
5
6
7
8
Scenario 3:
Quaternary
divergence
LGM
LIG
9
10
11
12
Scenario 4:
Neogene
divergence
LGM
LIG
13
14
15
16

Figure 2



23 **Figure 3**

24



25 **Figure 4**

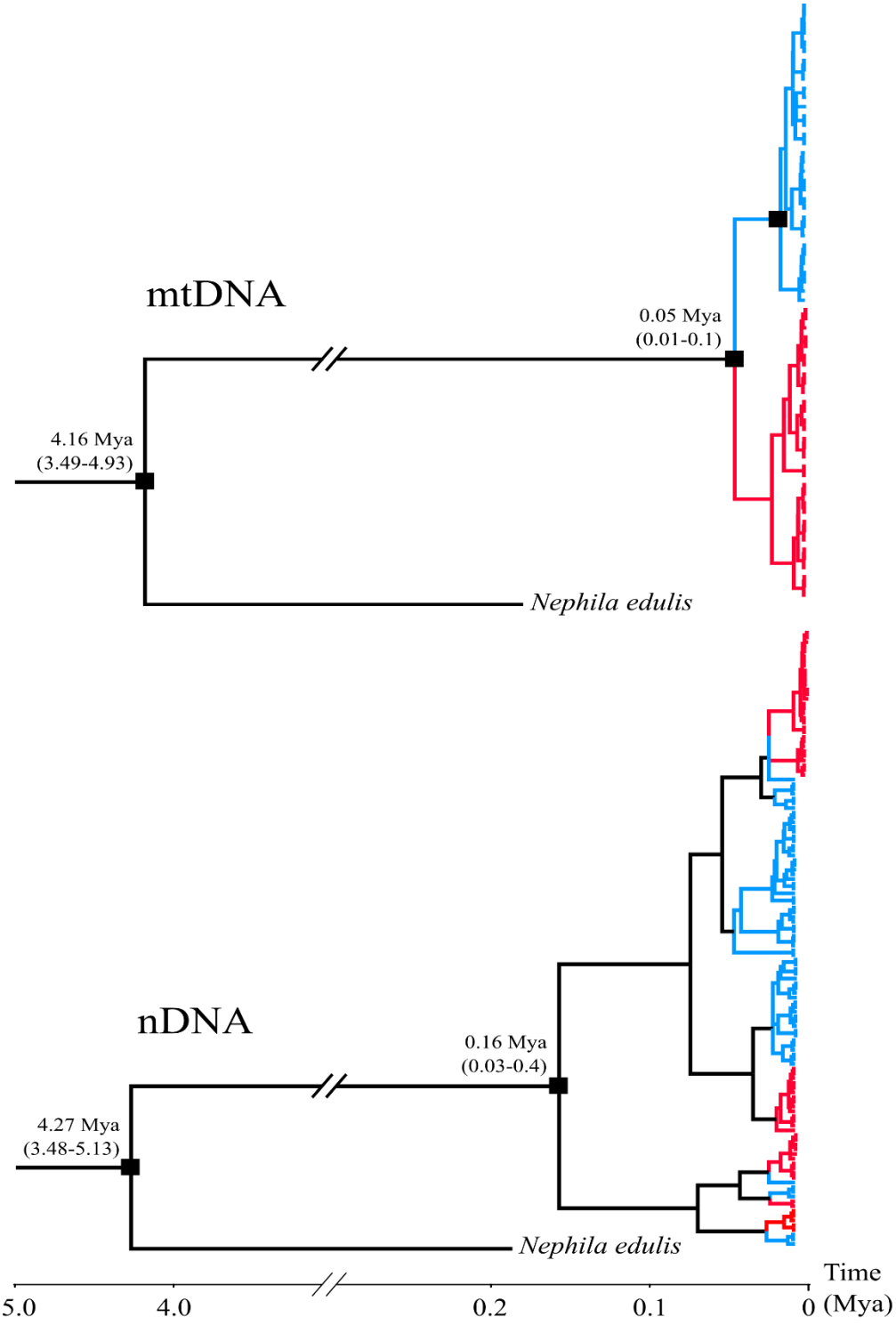

mtDNA
0.05 Mya
(0.01-0.1)
4.16 Mya
(3.49-4.93)
Nephila edulis
nDNA
0.16 Mya
(0.03-0.4)
4.27 Mya
(3.48-5.13)
Nephila edulis
Time
(Mya)
5.0
4.0
0.2
0.1
0

Figure 5

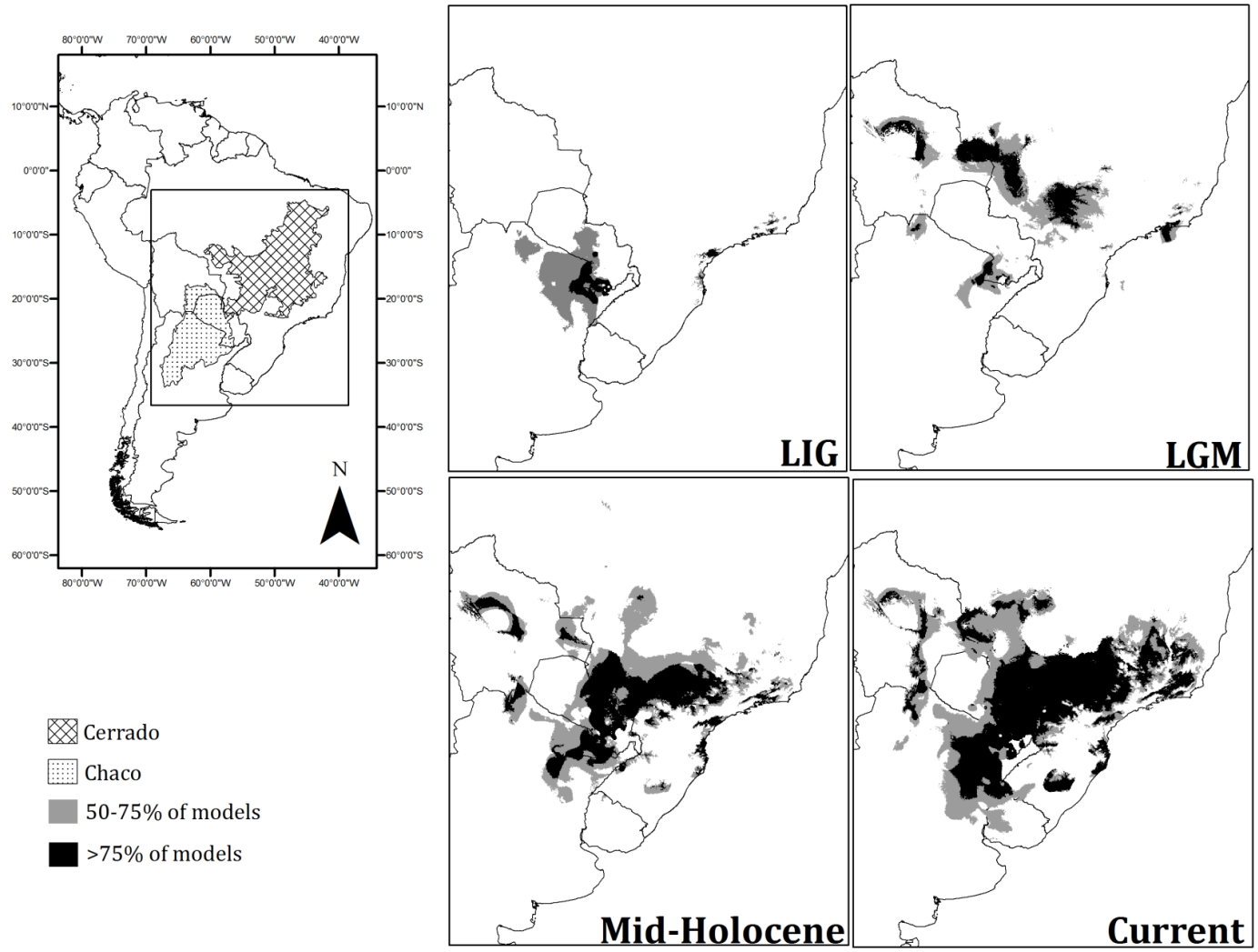

80°0'0"W
70°0'0"W
60°0'0"W
50°0'0"W
40°0'0"W
10°0'0"N
0°0'0"
10°0'0"S
20°0'0"S
30°0'0"S
40°0'0"S
50°0'0"S
60°0'0"S
N
Cerrado
Chaco
50-75% of models
>75% of models
LIG
LGM
Mid-Holocene
Current

Figure 6

Journal of Biogeography

SUPPORTING INFORMATION

Phylogeography of the dry vegetation endemic species *Nephila sexpunctata* (Araneae: Araneidae) suggests recent expansion of the Neotropical Dry Diagonal

Luiz Filipe de M. Bartoleti; Elen A. Peres; Thadeu Sobral-Souza; Fernanda von Hertwig M. Fontes; Márcio J. da Silva and Vera N. Solferini

APPENDIX S1. DETAILS ON METHODOLOGY.

Ecological Niche Modelling

Occurrence points were all considered as single occurrences in a 2.5' arc-resolution (5x5 km grid cell resolution at Equator region). The models were run with South America as background to consider the whole available area to species dispersion throughout its evolutionary history (Barve *et al*, 2011).

We used the 19 variables available on the WorldClim Database (Hijmans *et al*, 2005) and performed variables selection through a Jackkniffe procedure, using Maxent (Phillips & Dudik, 2008). Temperature Seasonality, Minimum Temperature of Coldest Month, Mean Temperature of Wettest Quarter and Precipitation of Warmest Quarter were the variables with highest contributions to model construction (Table S2) and therefore were selected for model calibration.

To infer potential species distribution, we used five algorithms that represent different modelling techniques (envelope, distance and background): (1) Bioclim (Nix, 1986); (2) Mahalanobis Distance (Farber & Kadmon, 2003); (3) Domain (Gower distance; Carpenter *et al*. 1993). (4) Support Vector Machines (SVM) (Tax & Duin, 2004) and (5) Maximum Entropy (Phillips & Dudik, 2008). For each algorithm we used 75-25% of occurrence points for training/testing. We evaluated the models using the TSS value (acceptable models present TSS values >0.5 - Allouche *et al*, 2006). We repeated this procedure 20 times for each algorithm and used the Lowest Present Threshold values (Pearson *et al*, 2007) to transform each map in binary. We then used the ensemble approach (Araújo & New, 2007) for computing the frequency of prediction by grid cell to create a consensus map for each algorithm. The last step was to concatenate all consensus maps and then obtain a consensus map for each climatic scenario.

Model Testing

Scenario 1 comprised four models:

a) A single panmictic population that presented no demographic variation over time, therefore with a constant size; b) a panmictic population with constant size that experienced exponential growth; c) a panmictic population that started to grow exponentially in the LIG and then experienced a strong bottleneck event in the LGM that extends to the present, as expected by the Pleistocene Refugia Model; d) a panmictic population that experienced a strong bottleneck event from LIG to LGM, and then displayed post-glacial exponential growth.

Scenario 2 represented two populations, one in Chaco and the other in Cerrado, with a divergence time estimated after the LGM (0.0-0.021Ma). In this scenario, the demographic events occurred before the populations diverge; therefore they experienced demographic fluctuations as a single population. The four models within this scenario correspond to the changes in population size before the split: a) a constant population; b) exponential growth of a previous constant-size population until the divergence; c) growth during the glacial with an interglacial bottleneck; d) glacial bottleneck with a post-glacial exponential growth.

Scenario 3 also represents two populations, but the divergence time between them was estimated in the Quaternary, before the LIG (0.120-2.588Ma). This scenario comprised four models similar to those from scenarios 1 and 2: a) constant populations; b) exponential growth of previous constant-size populations; c) exponential growth during the glacial with an interglacial bottleneck; d) glacial bottleneck with a post-glacial exponential growth. We considered that each population (Chaco and Cerrado) experienced demographic events independently, so each one had independent values for parameters like GrowthRate and population size after the bottleneck event, for instance.

For scenario 4, we also considered two populations but with a split estimated in the late Neogene (Pliocene/Miocene 2.588-23.03Ma). The models tested and the approach used were the same as for scenarios 2 and 3.

For scenarios 2, 3 and 4, after selecting the best fitting model among the four models, we contrasted the chosen one with a similar model that also considered recent bidirectional migration between the two populations. The model selected between this two were considered as the best fitting model for this scenario.

Scenario	Model	Nº populations	Demography	Migration?
1	1	Panmictic population.	No demographic events.	Not tested.
	2		Demographic expansion of a previous constant-size population.	
	3		Glacial expansion and interglacial retraction.	
	4		Glacial retraction and post-LGM expansion.	
2	5	Two populations with a post-LGM divergence.	No demographic events.	The best among the four models was contrasted with a similar model that considered bidirectional migration.
	6		Demographic expansion of a previous constant-size population.	
	7		Glacial expansion and interglacial retraction.	
	8		Glacial retraction and post-LGM expansion.	
3	9	Two populations with a Quaternary divergence.	No demographic events.	The best among the four models was contrasted with a similar model that considered bidirectional migration.
	10		Demographic expansion of a previous constant-size population.	
	11		Glacial expansion and interglacial retraction.	
	12		Glacial retraction and post-LGM expansion.	
4	13	Two populations with a late Neogene (Pliocene/ Miocene) divergence.	No demographic events.	The best among the four models was contrasted with a similar model that considered bidirectional migration.
	14		Demographic expansion of a previous constant-size population.	
	15		Glacial expansion and interglacial retraction.	
	16		Glacial retraction and post-LGM expansion.	

We developed custom Python scripts in order to simulate data on *ms* (Hudson, 2002), and performed 400,000 data simulations per scenario using the same number of individuals as the empirical data, which were divided in two groups according to their geographical region (Cerrado or Chaco). For Scenario 1, in order to simulate a single panmictic population, we used the same two groups as in the others scenarios, but with a very recent divergence time (close to present-day). This approach allowed us to simulate a panmictic population and use the same summary statistics in all scenarios.

We initially set all of the parameters (including θ for each marker) with flat prior uniform distributions in order to provide a wide range of prior sampling. Even though this approach increases computational work it is a safer exploratory step when there are uncertainties concerning prior distribution. We estimated times of the demographic events

based on the formula τ = Number of generations/4*N_e. In this case, the generation time for *N. sexpunctata* was adopted as 1 year, and N_e was calculated through the formula $N_e = \theta/\mu$ (for the mitochondrial marker), with θ being estimated on DnaSP and for μ we considered the rate estimated by Kuntner *et al* (2013) of 0.0112 substitutions/site/million years. We used this estimation of N_e for both mitochondrial and nuclear datasets due to the lack of information on the substitution rate of the nuclear marker.

Summary statistics of the simulated dataset were calculated using a PERL script written by Naoki Takebayashi (available at http://raven.iab.alaska.edu/~ntakebay/teaching/programming/coalsim/scripts/msSS.pl), [total nucleotide diversity, number of segregating sites, Tajima's D, nucleotide diversity within and between populations (π, ss, D, π_w, π_b, respectively)]. We grouped these statistics in vectors in order to find the most informative set of summary statistics (i.e., which vector more accurately identifies the best model given the simulated data). We performed a rejection step with 10 simulations for each model and used them as PODs (pseudo-observed data). The best vector of summary statistics would enhance the probability of choosing the true model over the average probability of choosing any other model (Tsai & Carstens, 2013).

After choosing the best set of summary statistics, we performed a parameter restriction in order to narrow their distributions and get a more reliable model choice. The restriction was made with the "abc" package in R with a threshold of 0.1% of the simulations. We then took the posterior distribution for each parameter of these exploratory runs and used them as priors for 1,600,000 new simulations comprising all the scenarios.

Finally, we used these simulations to perform the model choice. We used *msReject* (Ross-Ibarra *et al*, 2008) to perform a simple rejection with a threshold of 0.01% and calculate posterior probabilities of the competing models to find the best model within each scenario. We then gathered the three best models, performed new simulations and once again applied a simple rejection step in order to find the best overall model. Parameter estimation for all models are displayed in Figures S3-S6.

References:

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810–1819.

Carpenter, G., Gillison, A.N. & Winter, J. (1993) DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, **2**, 667–680.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hudson, R. R. (2002) Generating samples under a Wright-Fisher neutral model. *Bioinformatics*, **18**, 337-338.

Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.

Kuntner, M., Arnedo, M.A., Trontelj, P., Lokovsek, T., Agnarsson, I. (2013) A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution*, **69**, 961-979.

Nix, H.A. (1986) A biogeographic analysis of Australian elapid snakes. Atlas of Australian elapid snakes (ed. by R. Longmore), pp. 4–15. Australian Flora and Fauna Series 7, Australian Government Publishing Service,Canberra.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.

Phillips, S.J., Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

Ross-Ibarra, J., Wright, S.I., Foxe, J.P., Kawabe, A., DeRose-Wilson, L., Gos, G., Charlesworth, D., Gaut, B.S. (2008) Patterns of Polymorphism and Demographic History in Natural Populations of Arabidopsis lyrata. *PLoS ONE*, **3**(6): e2411

Tax, D.M.J. & Duin, R.P.W. (2004) Support Vector Data Description. *Machine Learning*, **54**, 45-66.

Tsai, Y-H. E., Carstens, B.C. (2013) Assessing model fit in phylogeographic investigations: An example from the North American willow *Salix melanopsis*. *Journal of Biogeography*, **40**, 131-141.

APPENDIX S2. SUPPLEMENTARY TABLES

Table S1: Occurrence points used for the Ecological Niche Modelling of *Nephila sexpunctata*. Here we grouped sampling points used in this work as well as occurrence points obtained on specieslink and GBIF. AR = Argentina, BR = Brazil, PAR = Paraguay.

Locality	Latitude	Longitude	Locality	Latitude	Longitude
AR - Paraná	-31.72559	-60.29947	BR -Tupã	-21.917172	-50.501272
AR - Federal	-30.93362	-58.72915	BR - São José do Rio Pardo	-21.595556	-46.888611
AR - Mercedes	-29.09571	-58.32959	BR - Campo Grande	-20.509839	-53.678824
AR - Santa Maria	-28.04155	-58.1006	BR - Ribas do Rio Pardo	-21.311625	-54.615461
AR – PN del Chaco	-26.80864	-59.60598	BR - Rancharia	-22.229167	-50.893056
AR - Basail	-27.86713	-59.28944	BR - Presidente Epitácio	-21.792760	-52.135545
AR - Aguaray	-22.242213	-63.736972	BR - Ribeirão Preto	-21.177500	-47.810278
AR - Corrientes	-28.681721	-57.858095	BR - Ibitinga	-21.757778	-48.828889
BR - Pirenópolis	-15.923383	-49.21405	BR - Lins	-21.678611	-49.7425
BR - Quaraí	-30.2072	-57.5547	BR - Bauru	-22.314722	-49.060556
BR - Cuiabá	-15.5961	-56.0967	BR - Três Lagoas	-20.751111	-51.678333
BR - Catalão	-18.113889	-47.621111	PAR - Villeta	-22.494	-58.012

Table S2: The Jackknife analysis results for 19 Worldclim variables and their respective contribution to species occurrence. Variables in bold are those with highest contributions to model construction.

Variable	Percent contribution	Permutation importance
Temperature Seasonality	37.2	26.5
Mean Temperature of Wettest Quarter	17.1	16.1
Min Temperature of Coldest Month	15.1	0.6
Precipitation of Warmest Quarter	13.6	7.1
Precipitation of Coldest Quarter	7.8	19.6
Precipitation of Wettest Quarter	3.6	7.8
Precipitation Seasonality	2.7	0
Mean Temperature of Coldest Quarter	2	14.5
Precipitation of Driest Quarter	0.4	7.3
Annual Mean Temperature	0.4	0.6
Temperature Annual Range	0	0
Annual Precipitation	0	0
Precipitation of Driest Month	0	0
Precipitation of Wettest Month	0	0
Isothermality	0	0
Max Temperature of Warmest Month	0	0
Mean Temperature of Driest Quarter	0	0
Mean Temperature of Warmest Quarter	0	0
Mean Diurnal Range	0	0

Table S3: TSS mean values for each algorithm used in the ENM for *Nephila sexpunctata*. sd = standard deviation.

Algorithm	TSS mean value (sd)
BIOCLIM	0.5 (0.254)
GOWER	0.683 (0.161)
MAHALANOBIS DISTANCE	0.742 (0.148)
MAXENT	0.658 (0.213)
SVM	0.8 (0.167)

Table S4: Pairwise F_{ST} values between *Nephila sexpunctata* populations. Upper/right: Nuclear dataset. Lower/left: Mitochondrial dataset. * $p < 0.05$.

	TP	**PI**	**CT**	**CG**	**PA**	**FE**	**ME**	**SM**	**CH**	**BA**
TP	-	0.009	0.018	0.084	0.337*	-0.024	0.003	-0.008	-0.036	0.11515
PI	0.190	-	0.086	-0.041	0.413*	0.109	-0.016	0.110	-0.059	0.26222*
CT	0.098	-0.088	-	0.028	-0.023	-0.023	-0.013	-0.078	0.012	-0.03849
CG	0.171	0.069	-0.091	-	0.337*	0.176*	-0.052	0.155*	-0.022	0.3762*
PA	0.415	0.645*	1.000	0.613*	-	0.228	0.232*	0.162	0.339*	0.11545
FE	0.738*	0.690*	0.833*	0.537*	0.910*	-	0.063*	-0.049	0.039	0.00667
ME	0.143	0.442*	0.415*	0.336	-0.043	0.659*	-	0.057	-0.026	0.16667*
SM	0.805*	0.753*	1.000*	0.655*	1.000*	0.000	0.714*	-	0.034	-0.02667
CH	0.460*	0.450*	0.382	0.204	0.571*	0.072*	0.392*	0.089*	-	0.17383*
BA	0.732*	0.696*	0.757*	0.568*	0.846*	0.476*	0.672*	0.571*	0.058	-

Table S5: Parameter estimation for the two overall highest supported models in the ABC analysis. In parenthesis, the minimum and maximum values of the posterior distribution.

Parameter	Model 4 (panmixia)	Model 6 (post-LGM divergence)
Θ COI	4.302 (0.8-13.406)	2.138 (0.472-7.228)
Θ CHP2	2.17 (0.426-5.713)	0.913 (0.36-2.59)
Growth Rate	30.06 (0.0-92.721)	3.503 (0.372-66.719)
Size before bottleneck (compared to Ne)	8.356 (7.025-9.035)	-
Divergence Time (in million years)	-	0.018 (0.012-0.021)
Expansion Time (in million years)	-	0.082 (0.012-0.117)

APPENDIX S3. SUPPLEMENTARY FIGURES

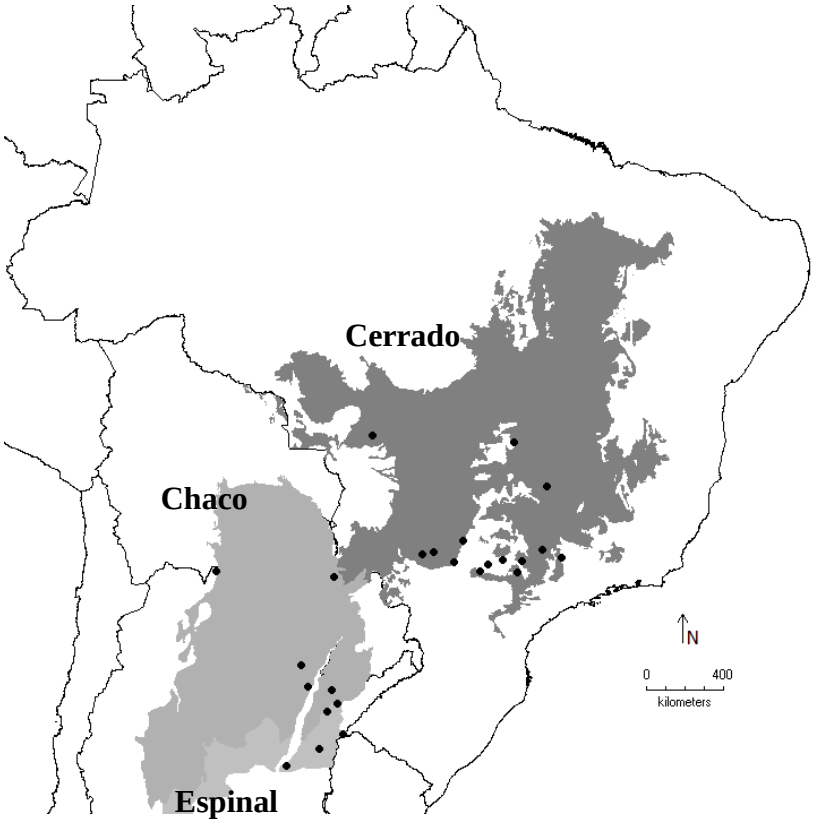


Figure S1: Map showing the known distribution of *N. sexpunctata* along Cerrado, Chaco and Espinal.

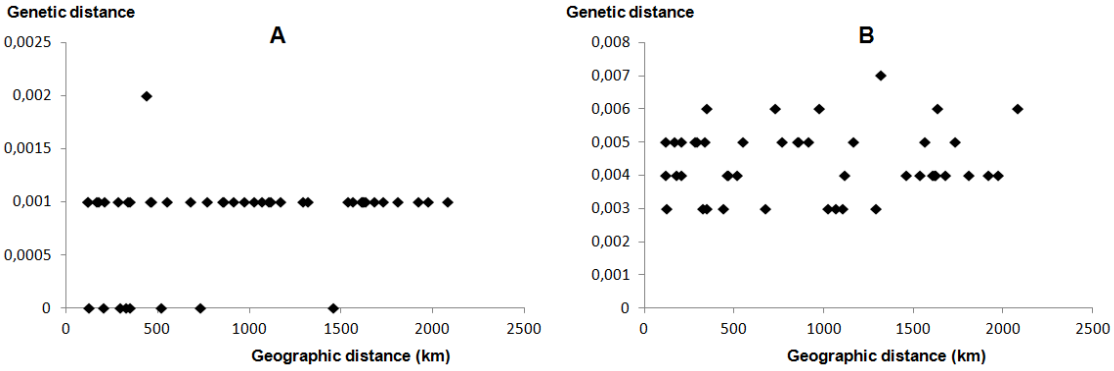


Figure S2: Genetic distance vs Geographic distance (km) graphs for *Nephila sexpunctata* populations. (a) Comparison using only the mitochondrial dataset (p = 0.038). (b) Comparison using the nuclear dataset (p = 0.319).

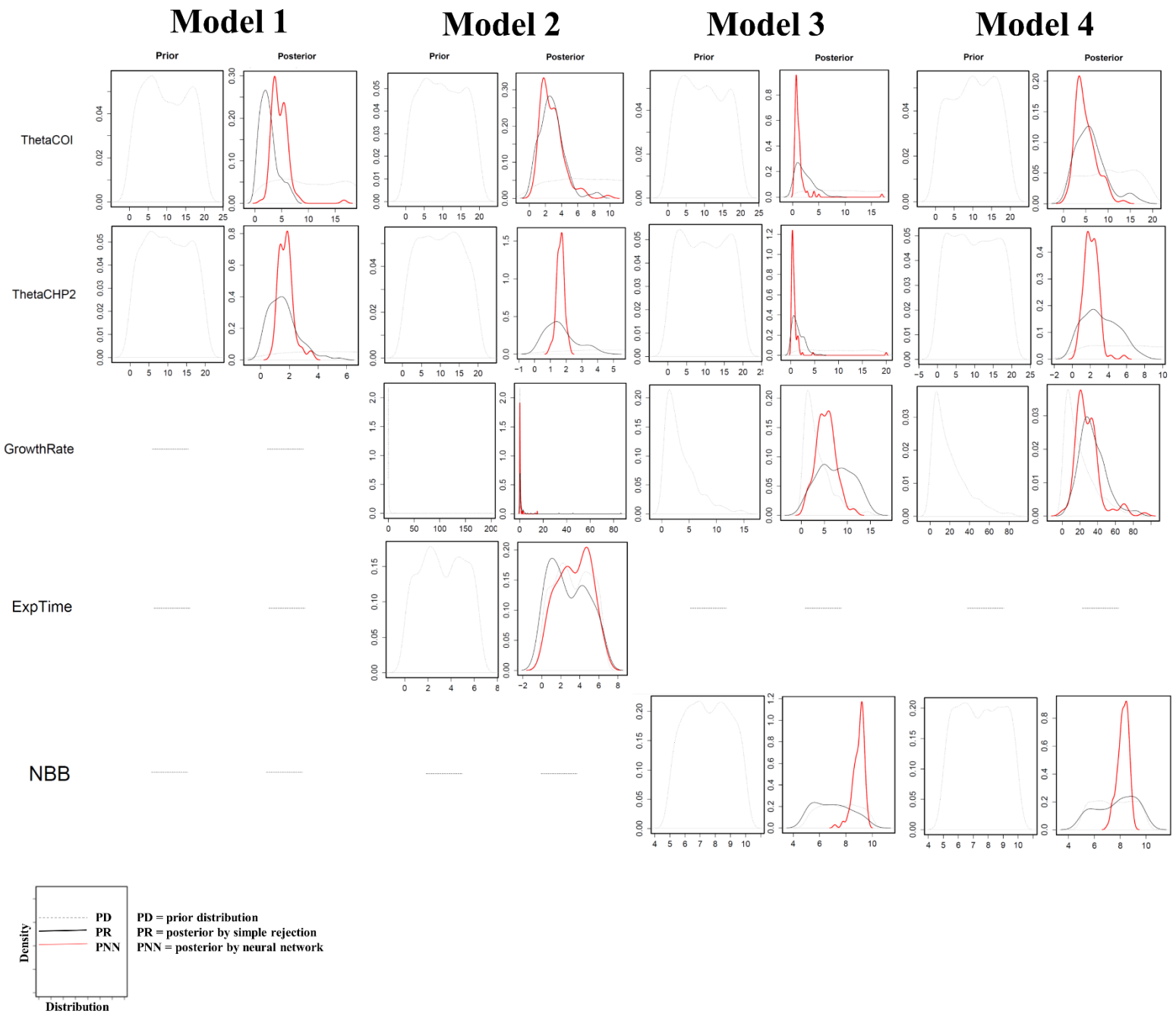


Figure S3: Prior and posterior distributions for each parameter within each model from Scenario 1 in the ABC analysis for *N. sexpunctata*. Absence of a chart in any model means that the parameter was not considered in that model. For each model, the left chart represents the prior distribution of the parameter (dashed line), while the right chart shows the posterior distribution by simple rejection (black line) and by neural network (red line), the last one was considered in this study.

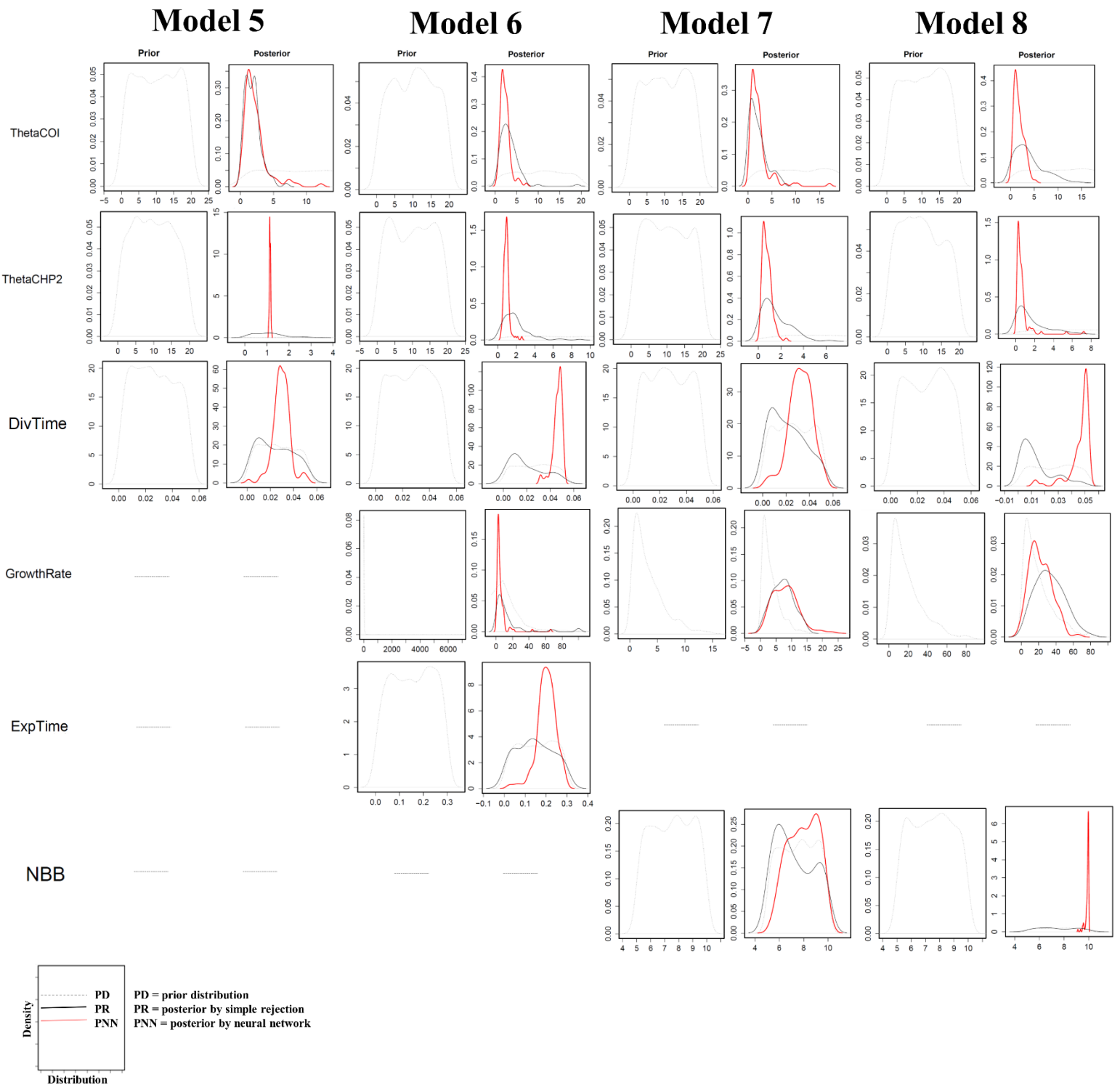


Figure S4: Prior and posterior distributions for each parameter within each model from Scenario 2 in the ABC analysis for *N. sexpunctata*. Absence of a chart in any model means that the parameter was not considered in that model. For each model, the left chart represents the prior distribution of the parameter (dashed line), while the right chart shows the posterior distribution by simple rejection (black line) and by neural network (red line), the last one was considered in this study.

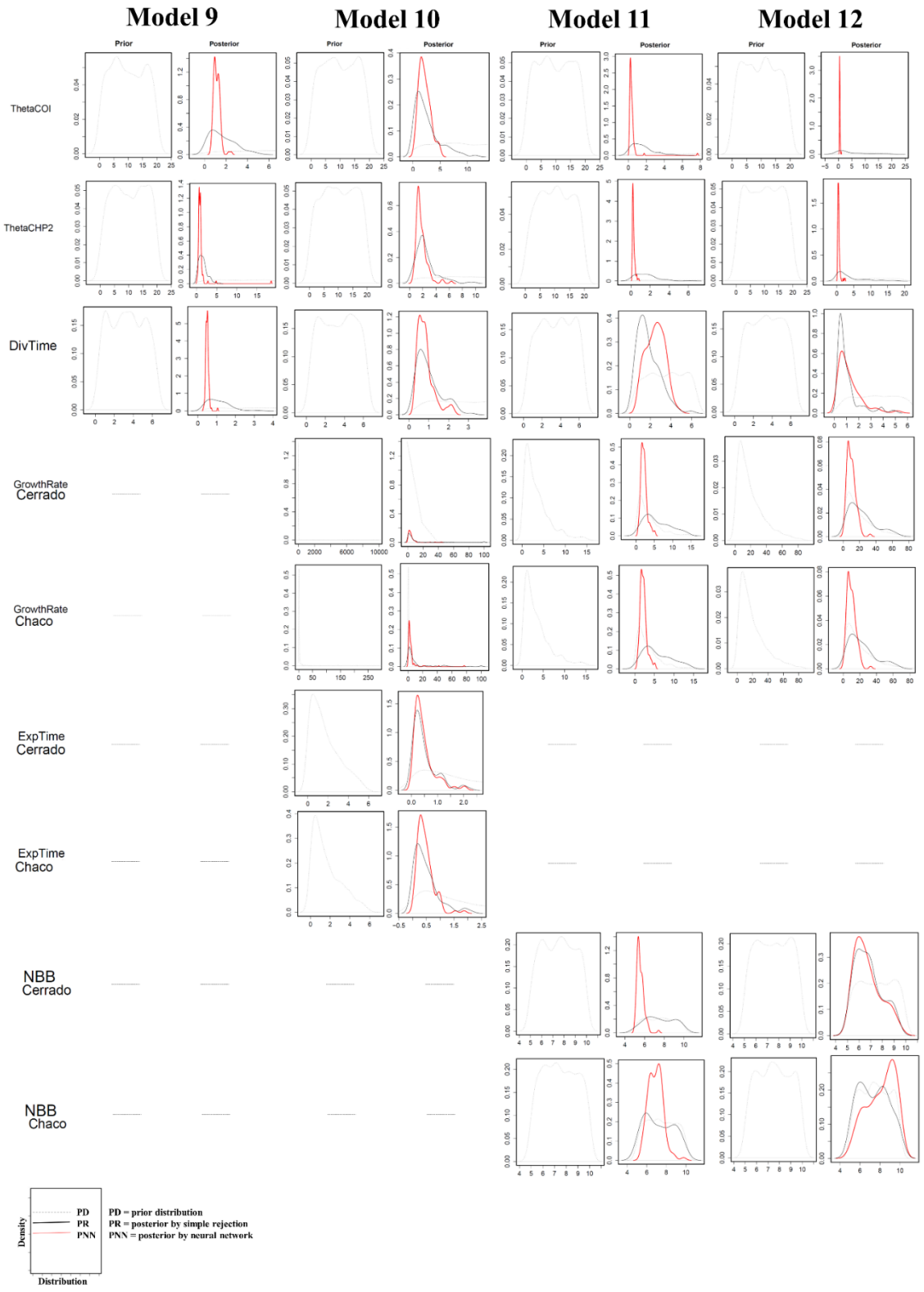


Figure S5: Prior and posterior distributions for each parameter within each model from Scenario 3 in the ABC analysis for *N. sexpunctata*. Absence of a chart in any model means that the parameter was not considered in that model. For each model, the left chart represents the prior distribution of the parameter (dashed line), while the right chart shows the posterior distribution by simple rejection (black line) and by neural network (red line), the last one was considered in this study.

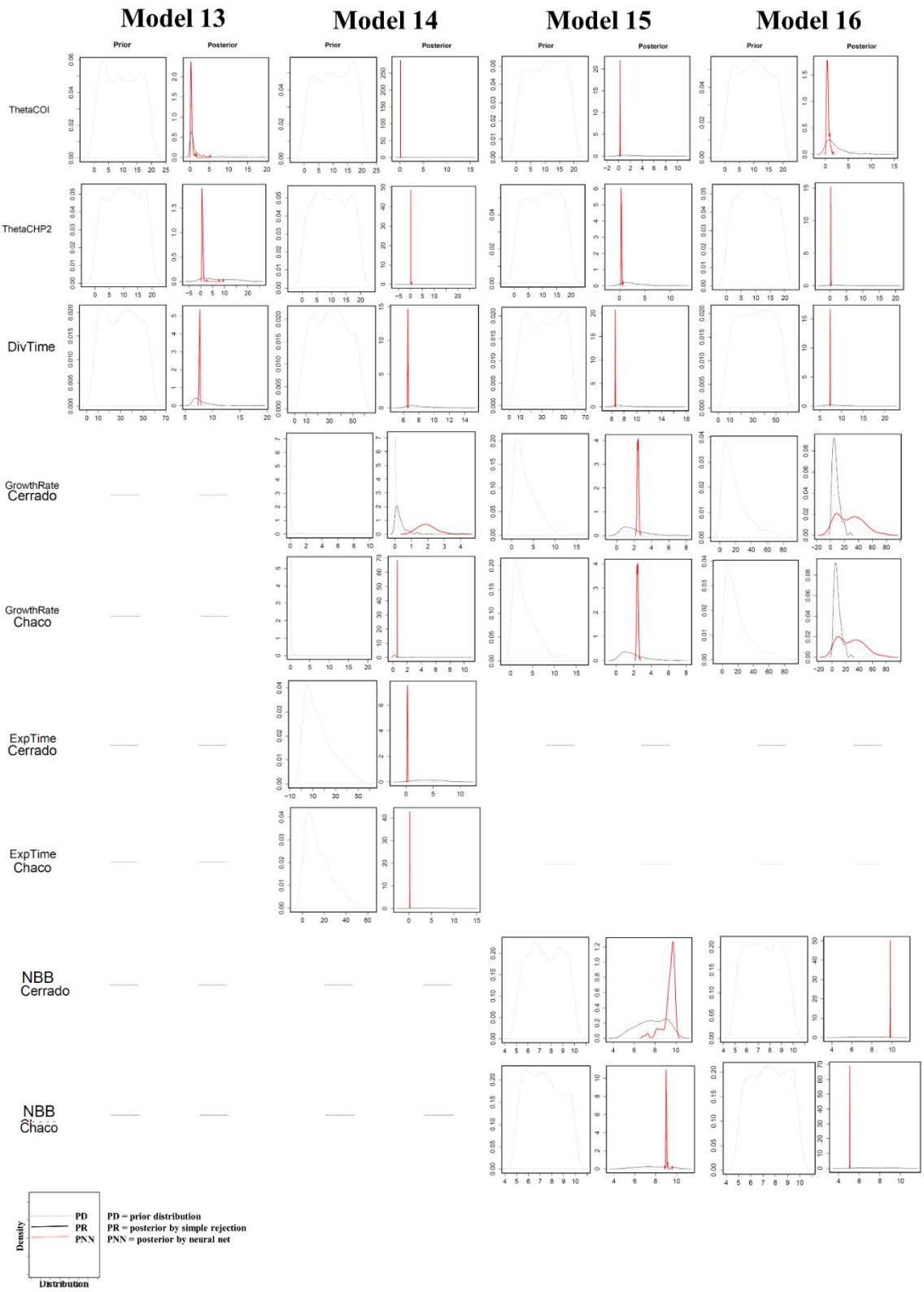


Figure S6: Prior and posterior distributions for each parameter within each model from Scenario 4 in the ABC analysis for *N. sexpunctata*. Absence of a chart in any model means that the parameter was not considered in that model. For each model, the left chart represents the prior distribution of the parameter (dashed line), while the right chart shows the posterior distribution by simple rejection (black line) and by neural network (red line), the last one was considered in this study.

DISCUSSÃO GERAL

Padrões filogeográficos das espécies de *Nephila* na América do Sul

As duas espécies de *Nephila* puderam ser utilizadas com sucesso no presente estudo pois apresentaram sinais filogeográficos adequados para a inferência de eventos demográficos. Em ambos os casos, foram encontradas linhagens mitocondriais geograficamente informativas, enquanto os marcadores nucleares apresentaram estruturação genética consideravelmente mais baixa. A estruturação mais pronunciada no DNA mitocondrial, provavelmente, é devido à diferença de tamanho efetivo entre o genoma da mitocôndria e do núcleo (Moore, 1995), permitindo uma visualização mais clara da história evolutiva por meio dos marcadores mitocondriais (Zink & Barrowclough, 2008). Além disso, fatores como introgressão do genoma mitocondrial, a separação incompleta de linhagens, e a dispersão diferencial entre machos e fêmeas, também podem contribuir para os padrões observados (Toews & Brelsford, 2012).

A alta estruturação genética encontrada nas espécies Neotropicais contrasta com estudos prévios realizados com *Nephila* em outros continentes, que evidenciam boa manutenção de fluxo gênico mesmo em populações separadas por amplas distâncias ou por acidentes geográficos como cadeias montanhosas (Tso *et al*, 2002; Lee *et al*, 2004; Su *et al*, 2007; Su *et al*, 2011). Nossas análises indicam que, apesar da alta estruturação, *N. clavipes* e *N. sexpuctata* também apresentam grande capacidade de dispersão, pois diversos locais de amostragem apresentam mais de uma linhagem genética, indicando contato secundário entre grupos previamente isolados. Além disso, os testes de modelos ressaltaram a importância da dispersão nessas espécies, uma vez que modelos panmíticos ou que consideravam migração entre as populações foram os mais prováveis para ambas as espécies. Enquanto os estudos com espécies asiáticas utilizaram populações separadas, em sua maioria, por uma matriz oceânica numa dinâmica de ilhas, as espécies Neotropicais são separadas por diferentes biomas, que apresentam uma dinâmica complexa e ainda não muito bem elucidada. O fato de as espécies Neotropicais apresentarem estruturação e divergência genéticas maiores que as espécies asiáticas realça o importante papel das flutuações climáticas do Quaternário em promover isolamento e diversificação na região.

Os níveis de variabilidade genética nas duas espécies foram bastante discrepantes: para *N. sexpunctata*, a variabilidade dos marcadores mitocondriais foi muito mais baixa que o encontrado em trabalhos com outras espécies de *Nephila* que também utilizaram sequências de COI (Lee *et al*, 2004; Su *et al*, 2007; Su *et al*, 2011); já para *N. clavipes*, a variabilidade

mitocondrial foi alta, similar aos demais trabalhos. No caso de *N. sexpunctata*, a baixa variabilidade parece ser associada a eventos climáticos que restringiram a distribuição da espécie próximo ao último glacial (~21 mil anos atrás), que acarretaram em gargalos populacionais e consequente diminuição da variabilidade genética (Bonatelli *et al*, 2014). Já *N. clavipes*, que possui mais ampla distribuição e maior tolerância ambiental, apresentou estabilidade demográfica ao longo do último glacial, sem perda de variabilidade genética.

Ambas espécies divergiram dos *outgroups* utilizados nas análises (*N. inaurata* para *N. clavipes*, e *N. edulis* para *N. sexpunctata*) entre o Mioceno e o Plioceno. As estimativas de divergências são concordantes com o trabalho de Kuntner *et al*, 2013, que reconstruiu a filogenia do gênero e propôs que as Américas teriam sido colonizadas independentemente duas vezes – uma por uma linhagem oriunda da África (da qual fazem parte *N. inaurata* e *N. clavipes*) e por uma linhagem da Ásia/Oceania (que originou *N. edulis* e *N. sexpunctata*). As divergências relativamente recentes (<10 milhões de anos para ambos os casos), quando a América já se encontrava separada dos demais continentes, reforçam a ideia de colonização. Estudos com plantas (e.g. Pennington & Dick, 2004; Renner, 2004; Christenhusz & Chase, 2013) e animais (e.g. Carranza & Arnold, 2003; Blaimer, 2012) demonstraram que, ao contrário do que se imaginava, a biota sul-americana não se manteve isolada no período entre a separação do continente africano e o fechamento do istmo do Panamá, e que movimentações transoceânicas foram recorrentes para diversos grupos e explicam boa parte da diversidade existente no continente.

História biogeográfica recente da América do Sul revelada pela filogeografia de *N. clavipes* e *N. sexpunctata*

A análise filogeográfica das duas espécies de *Nephila* permitiram a realização de interessantes inferências sobre a história recente dos biomas sul-americanos. No caso de *N. clavipes*, pode-se notar que tantos eventos geológicos quanto climáticos foram responsáveis por promover diversificação nos organismos que ocorrem em florestas úmidas. O soerguimento dos Andes colombianos, principalmente da Cordilheira Central (entre o Mioceno e o Plioceno) parece ter isolado diferentes linhagens de *N. clavipes* que desde então evoluem independentemente, como indicado pela reconstrução filogenética da espécie e pelo teste de modelos. Esse padrão também foi encontrado em outros organismos (e.g. Muñoz-Ortiz *et al*, 2014), reforçando o soerguimento da Cordilheira Central como um importante evento vicariante. Além desse evento geológico, as flutuações climáticas do Pleistoceno parecem ter

ocasionado diversificação nos biomas brasileiros, com linhagens relacionadas à Amazônia, Cerrado e Mata Atlântica. Essas linhagens, entretanto, não se mantiveram isoladas em apenas um bioma, e uma possível rota de conexão entre Amazônia e Mata Atlântica durante o Último Máximo Glacial pelo centro do Cerrado foi corroborada pelos nossos dados genéticos e de teste de modelos, com o Cerrado sendo a principal fonte de migrantes para ambas florestas. Quanto à Mata Atlântica, as partes Norte e Sul (separadas por diferentes regimes climáticos) apresentaram grande similaridade genética, provavelmente ocasionada por boa manutenção do fluxo gênico entre as duas partes, ao contrário do que é encontrado para muitos organismos (e.g. Prates *et al*, 2016; Peres *et al*, 2017).

Para *N. sexpunctata*, a recente coalescência de todas as sequências indica o grande papel dos eventos climáticos do Quaternário na diversificação da espécie. O Último Máximo Glacial parece ter mantido as condições climáticas demasiadamente frias e secas, mesmo para organismos que ocorrem em formações abertas; dessa maneira, *N. sexpunctata* parece ter sofrido uma diminuição na sua área de ocorrência nesse período, o que levou à baixa variabilidade genética encontrada na espécie – também observada em outros organismos que ocorrem nessa região (Babb *et al*, 2011; Brito *et al*, 2013). A expansão demográfica pós-glacial detectada pelos dados genéticos e corroborada pelo teste de modelos também concorda com a expansão da área de ocorrência inferida pela modelagem de distribuição. Assim, as flutuações climáticas do Quaternário parecem ter sido responsáveis por eventos demográficos que afetaram as espécies que vivem na Diagonal Seca.

CONSIDERAÇÕES FINAIS

Os trabalhos aqui apresentados corroboram pontos importantes sobre a diversificação da região Neotropical, ressaltando que eventos geológicos e climáticos foram responsáveis por esse processo em diversos tipos de vegetação no continente. Por outro lado, foram observadas particularidades em cada espécie que contrastaram com hipóteses previamente estabelecidas, o que indica que diferentes espécies podem ter respondido de maneiras distintas a esses eventos. Assim como observado em outros continentes, as espécies de *Nephila* da América do Sul constituíram adequados modelos filogeográficos, e o estudo de aranhas e outros invertebrados deve ser encorajado. A utilização de espécies amplamente distribuídas mostrou-se fundamental para a melhor compreensão da história evolutiva de regiões altamente diversas. Por fim, nossos dados corroboram que os biomas da região Neotropical possuem uma história complexa e intricada, e o aumento no número de estudos filogeográficos certamente colaborará para a melhor compreensão da biogeografia história da América do Sul.

REFERÊNCIAS BIBLIOGRÁFICAS

Ab'Saber AN (1977) Os domínios morfoclimáticos na América do Sul. Primeira aproximação. Geomorfologia, 53, 1–23.

Álvarez-Padilla F, Dimitrov D, Giribet G, Hormiga G (2009) Phylogenetic relationships of the spider family Tetragnathidae (Araneae, Araneoidea) based on morphological and DNA sequence data. *Cladistics*, **25(2)**, 109–146.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36(1)**, 3–15.

Babb PL, Fernandez-Duque E, Baiduc CA, Gagneux P, Evans S, Schurr TG (2011) mtDNA diversity in Azara's Owl Monkeys (*Aotus azarai azarai*) of the Argentinean Chaco. *American Journal of Physical Anthropology*, **146**, 209-224.

Batalha-filho H, Fjeldsa J, Fabre P, Miyaki CY (2013) Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. *Journal of Ornithololology*, **154**, 41–50.

Blaimer BB (2012) Acrobat ants go global – Origin, evolution and systematic of the genus Crematogaster (Hymenoptera:Formicidae). *Molecular Phylogenetics and Evolution*, **65(2)**, 421-436.

Beaumont MA (2010) Approximate Bayesian Computation in Evolution and Ecology. *Annual Reviews*, **41**, 379-406.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754–3774.

Bell JR, Bohan DA, Shaw EM, Weyman GS (2005) Ballooning dispersal using silk: world fauna, phylogenies, genetics and models. *Bulletin of entomological research*, **95**, 69–114.

Bidegaray-Batista L, Sánchez-Gracia A, Santulli G, Maiorano L, Guisan A, Vogler AP, Arnedo MA (2016) Imprints of multiple glacial refugia in the Pyrenees revealed by phylogeography and palaeodistribution modelling of an endemic spider. *Molecular Ecology*, **25(9)**, 2046–2064.

Blaimer BB (2012) Acrobat ants go global – Origin, evolution and systematic of the genus *Crematogaster* (Hymenoptera:Formicidae). *Molecular Phylogenetics and Evolution*, **65(2)**, 421-436.

Bonatelli IAS, Perez MF, Peterson AT, Taylor NP, Zappi DC, Machado MC, Koch I, Pires AHC, Moraes EM (2014) Interglacial microrefugia and diversification of a cactus species complex: phylogeography and palaeodistributional reconstructions for *Pilosocereus aurisetus* and allies. *Molecular Ecology*, **23**, 3044–3063.

Brito RM, Francisco FO, Françoso E, Santiago LR, Arias MC (2013) Very low mitochondrial variability in a stingless bee endemic to Cerrado. *Genetics and Molecular Biology*, **36(1)**, 124-128.

Carnaval AC, Moritz C (2008) Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, **35**, 1187–1201.

Carranza S, Arnold EN (2003) Investigating the origin of transoceanic distributions: mtDNA shows *Mabuya* lizards (Reptilia, Scincidae) crossed the Atlantic twice. *Systematics and Biodiversity*, **1(2)**, 275-282.

Carvalho AL, de Britto MR, Fernandes DS (2013) Biogeography of the lizard genus *Tropidurus* Wied-Neuwied, 1825 (Squamata:Tropiduridae): distribution, endemism, and area relationships in South America. *PLoS ONE*, **8**, e59736.

Carvalho CJB de, Almeida EAB (2011) Biogeografia da América do Sul: padrões & processos. Roca, São Paulo.

Christenhusz MJM, Chase MW (2013) Biogeographical patterns of plants in the Neotropics: dispersal rather than plate tectonics is most explanatory. *Botanical Journal of the Linnean Society*, **171**, 277–286.

Cooper SJB, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Molecular ecology*, **20(15)**, 3219–36.

Costa LP (2003) The historical bridge between the Amazon and the Atlantic Forest of Brazil: a study of molecular phylogeography with small mammals. *Journal of Biogeography*, **30**, 71–86.

Crews SC, Gillespie RG (2014) Desert salt flats as oases for the spider *Saltonia incerta* Banks (Araneae: Dictynidae). *Ecology and Evolution*, **4**, 3861–3874.

Croucher PJP, Oxford GS, Gillespie RG (2011) Population structure and dispersal in a patchy landscape: nuclear and mitochondrial markers reveal area effects in the spider *Theridion californicum* (Araneae: Theridiidae). *Biological Journal of the Linnean Society*, **104(3)**, 600–620.

Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, **25**, 410-418.

Dimitrov D, Benavides LR, Arnedo MA, Giribet G, Griswold CE, Scharff N, Hormiga G (2016) Rounding up the usual suspects: a standard target-gene approach for resolving the

interfamilial phylogenetic relationships of ecribellate orb-weaving spiders with a new family-rank classification (Araneae, Araneoidea). *Cladistics*, **33**, 221-250.

Garrick RC, Bonatelli IAS, Hyseni C, Morales A, Pelletier TA, Perez MF, Rice E, Satler JD, Symula RE, Thomé MTC, Carstens BC (2015) The evolution of phylogeographic data sets. *Molecular Ecology*, **24**, 1164–1171.

Haffer J (1969) Speciation in Amazonian forest birds. *Science*, **165**, 131-137.

Hodge MA, Uetz GW (1992) Antipredator benefits of single- and mixed-species grouping by *Nephila clavipes* (L.) (Araneae, Tetragnathidae). *The Journal of Arachnology*, **20**, 212–216.

Hoorn C, Wesselingh FP, Steege H *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science*, **330**, 927–931.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Knowles LL, Carstens BC, Keat ML (2007) Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, **17(11)**, 940-946.

Kuntner M (2006) Phylogenetic systematics of the Gondwanan nephilid spider lineage Clitaetrinae (Araneae, Nephilidae). *Zoologica Scripta*, **35(1)**, 19–62.

Kuntner M, Agnarsson I (2011) Phylogeography of a successful aerial disperser: the golden orb spider *Nephila* on Indian Ocean islands. *BMC Evolutionary Biology.* **11**,119.

Kuntner M, Arnedo MA, Trontelj P, Lokovsek T, Agnarsson I (2013) A molecular phylogeny of nephilid spiders: Evolutionary history of a model lineage. *Molecular Phylogenetics and Evolution*, **69**, 961-979.

Lee J, Jiang L, Su Y, Tso I (2004) Is Central Mountain Range a geographic barrier to the giant wood spider *Nephila pilipes* (Araneae: Tetragnathidae) in Taiwan? A population genetic approach. *Zoological Studies*, **43(1)**, 112–122.

Lee VMJ, Kuntner M, Li D (2015) Ballooning behavior in the golden orbweb spider *Nephila pilipes* (Araneae:Nephilidae). *Frontiers in ecology and evolution*, **3**, 2:1-5.

Lima-Ribeiro MS, Diniz-Filho JAF (2012) Modelando a distribuição geográfica das espécies no passado: uma abordagem promissora em paleoecologia. *Revista brasileira de paleontologia*, **15(3)**, 371-385.

Magalhaes IL, Oliveira U, Santos FR, Vidigal TH, Brescovit AD, Santos AJ (2014) Strong spatial structure, Pliocene diversification and cryptic diversity in the Neotropical dry forest spider *Sicarius cariri*. *Molecular Ecology*, **23**, 5323–5336.

Moore CW (1977) The life cycle, habitat and variation in selected web parameters in the spider *Nephila clavipes* Koch (Araneidae). *American Midland Naturalist*, **98(1)**, 95–108.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, **49**, 718–726.

Morley RJ (2000) Origin and evolution of tropical rain forests. John Wiley & Sons Ltd, Chichester.

Muñoz-Ortiz A, Velásquez-Álvarez AA, Guarnizo CE, Crawford A (2014) Of peaks and valleys: testing the roles of orogeny and habitat heterogeneity in driving allopatry in mid-elevation frogs (Aromobatidae: *Rheobates*) of the northern Andes. *Journal of Biogeography*, **42**, 193-205.

Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Pennington RT, Dick CW (2004) The role of immigrants in the assembly of the South American rainforest tree flora. *Philosophical Transactions of the Royal Society B*, **359**, 1611-1622.

Peres EA, Sobral-Souza T, Perez MF, Bonatelli IAS, Silva DP, Silva MJ, Solferini VN (2015) Pleistocene Niche Stability and Lineage Diversification in the Subtropical Spider *Araneus omnicolor* (Araneidae). *PLoS ONE*, **10(4)**, e0121543.

Peres EA, Silva MJ, Solferini VN (2017) Phylogeography of the spider *Araneus venatrix* (Araneidae) suggests past connections between Amazon and Atlantic rainforests. *Biological Journal of the Linnean Society* (in press). doi: https://doi.org/10.1093/biolinnean/blx036

Por FD (1992) Sooretama: the Atlantic rain forest of Brazil. The Hague: SPB Academic Publishing.

Potter PE, Szatmari P (2009) Global Miocene tectonics and the modern world. *Earth-Science Reviews*, **96**, 279–295.

Prates I, Rivera D, Rodrigues MT, Carnaval AC (2016) A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Journal of Biogeography*, **25**, 5174–5186.

Quiñones-Lebrón SG, Kralj-Fišer S, Gregorič M, Lokovšek T, Čandek K, Haddad CR, Kuntner M (2016) Potential costs of heterospecific sexual interactions in golden orbweb spiders (*Nephila* spp.). *Scientific Reports*, **6**, doi: 10.1038/srep36908

Renner S (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, **165:S4**, S23-S33.

Rittschof CC (2010) Male density affects large-male advantage in the golden silk spider, *Nephila clavipes*. *Behavioral Ecology*, **21(5)**, 979–985.

Robinson MH, Mirick H (1971) The predatory behavior of the golden-web spider *Nephila clavipes* (Araneae: Araneidae). *Psyche*, **78(3)**, 123–139.

Rull V (2011) Neotropical biodiversity: timing and potential drivers. *Trends in Ecology & Evolution,* **26,** 508–513.

Ruskin BG, Dávila FM, Hoke GD, Jordan TE, Astini RA, Alonso R (2011) Stable isotope composition of middle Miocene carbonates of the Frontal Cordillera and Sierras Pampeanas: did the Paranaense seaway flood western and central Argentina? *Palaeogeography, Palaeoclimatology. Palaeocology.* **308**, 293–303.

Rypstra AL (1985) Aggregations of *Nephila clavipes* (L.) (Araneae, Araneidae) in relation to prey availability. *Journal of Arachnology,* **13(1)**, 71–78.

Santos AJ, Brescovit AD (2001) Lycosidae - A revision of the South American spider genus. *Andrias,* **15,** 75–90.

Silva JMC (1995) Biogeographic analysis of the South American Cerrado avifauna. *Steenstrupia*, **21**, 49-67.

Simon MF, Grether R, de Queiroz LP, Skema C, Pennington RT, Hughes CE (2009) Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences USA*, **106**, 20359–20364.

Sobral-Souza T, Lima-Ribeiro MS, Solferini VN (2015). Biogeography of Neotropical rainforests: past connections between Amazon and Atlantic Forest detected by ecological niche modeling. *Evolutionary Ecology*, doi: 10.1007/s10682-015-9780-9.

Su Y-C, Chang Y-H, Lee S-C, Tso, I-M (2007) Phylogeography of the giant wood spider (*Nephila pilipes*, Araneae) from Asian-Australian regions. *Journal of Biogeography*, **34(1)**, 177–191.

Su Y-C, Chang Y-H, Smith D, Zhu M-S, Kuntner M, Tso I-M (2011) Biogeography and speciation patterns of the golden orb spider genus *Nephila* (Araneae: Nephilidae) in Asia. *Zoological science*, **28(1)**, 47–55.

Sunnaker M, Busetto AG, Numminem E, *et al*. (2013) Approximate Bayesian Computation. *PLoS Computational Biology*, doi: http://dx.doi.org/10.1371/journal.pcbi.1002803.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Tso I-M, Tai P-L, Ku T-H, Kuo C-H, Yang E-C (2002) Colour-associated foraging success and population genetic structure in a sit-and-wait predator *Nephila maculata* (Araneae: Tetragnathidae). *Animal Behaviour*, **63(1)**, 175–182.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Vanzolini PE, Williams EE (1981) The vanishing refuge: A mechanism for ecogeographic speciation. *Papéis avulsos de Zoologia*, **34(23)**, 251-255.

Winston ME, Kronauer DJC, Moreau CS (2017) Early and dynamic colonization of Central America drives speciation in Neotropical army ants. *Molecular Ecology*, **26**, 859–870.

World Spider Catalog (2017) World Spider Catalog. Natural History Museum Bern, online at http://wsc.nmbe.ch, version 18.0, accessed on {06.07.17)

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107–2121.

ANEXOS

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Filogeografia das espécies de Nephila (Araneae: Araneidae) da América do Sul**, não infringem os dispositivos da Lei n.° 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 16 de Agosto de 2017

Assinatura : Luiz F. de M. Bartoleti
Nome do(a) autor(a): **Luiz Filipe de Macedo Bartoleti**
RG n.° 30.710.327-4

Assinatura : _____
Nome do(a) orientador(a): **Vera Nisaka Solferini**
RG n.° 4763396

COORDENADORIA DE PÓS-GRADUAÇÃO
INSTITUTO DE BIOLOGIA
Universidade Estadual de Campinas
Caixa Postal 6109. 13083-970, Campinas, SP, Brasil
Fone (19) 3521-6378. email: cpgib@unicamp.br

UNICAMP

IB

DECLARAÇÃO

Em observância ao **§5º do Artigo 1º da Informação CCPG-UNICAMP/001/15**, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Tese de Doutorado, intitulada "***Filogeografia das espécies de Nephila (Araneae: Araneidae) da América do Sul***", desenvolvida no Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura: Luiz F. de M. Bartoleti
Nome do(a) aluno(a): Luiz Filipe de Macedo Bartoleti

Assinatura: Solferini
Nome do(a) orientador(a): Vera Nisaka Solferini

Data: 16 de Agosto de 2017